Approximations and Solution Estimates in Optimization

Johannes O. Royset
Operations Research Department
Naval Postgraduate School
joroyset@nps.edu

Abstract. Approximation is central to many optimization problems and the supporting theory provides insight as well as foundation for algorithms. In this paper, we lay out a broad framework for quantifying approximations by viewing finite- and infinite-dimensional constrained minimization problems as instances of extended real-valued lower semicontinuous functions defined on a general metric space. Since the Attouch-Wets distance between such functions quantifies epi-convergence, we are able to obtain estimates of optimal solutions and optimal values through estimates of that distance. In particular, we show that near-optimal and near-feasible solutions are effectively Lipschitz continuous with modulus one in this distance. We construct a general class of approximations of extended real-valued lower semicontinuous functions that can be made arbitrarily accurate and that involve only a finite number of parameters under additional assumptions on the underlying metric space.

Keywords: epi-convergence, Attouch-Wets distance, epi-splines, solution stability, approximation theory, near-optimality, near-feasibility, rate of convergence.

Date: April 6, 2016

1 Introduction

Solutions of many optimization problems are inaccessible by direct means and one is forced to settle for solutions of approximate problems. A central challenge is then to ensure that solutions of approximate problems are indeed approximate solutions of the original problems. Moreover, the degree of approximation becomes theoretically and practically important. The subject has been studied extensively; see, e.g., [17, 1, 23, 24] for foundations and [14, 25] for applications in machine learning and stochastic optimization. In this paper, we quantify the error in optimal values, optimal solutions, near-optimal solutions, and near-optimal near-feasible solutions for approximate problems defined on general metric spaces. In particular, we obtain a sharp Lipschitz-stability result for near-optimal solutions with a Lipschitz modulus of 1. We also construct a class of "elementary" functions called epi-splines that are given by a finite number of parameters, but still approximate to an arbitrary level of accuracy any extended real-valued lower semicontinuous (lsc) functions defined on a separable metric space. Since such lsc functions abstractly represent a large class of optimization problems, epi-splines therefore provide fundamental approximations of such problems.

The development relies heavily on set-convergence of epi-graphs, which goes back to the pioneering work of Wijsman [32, 33] and Mosco [22], and was coined *epi-convergence* by Wets [31]. This notion of convergence is the only natural choice for minimization problems as it guarantees the convergence of optimal solutions and optimal values of approximate problems to those of a limiting problem. Quantification of the distance between epi-graphs, which then leads to a quantification of epi-convergence, is placed on a firm footing in [4, 2, 5] with the development of the Attouch-Wets (aw) distance; see also [9, 11, 12, 10]. We follow these lines and especially those of [6, 7] that utilize such quantification as the basis for solution estimates in minimization problems. In contrast to these two papers, which deal with normed linear spaces, we consider general metric spaces. Also, our Lipschitz-stability result for near-optimal solutions goes beyond that of [7] as it does not require convexity, and we consider near-optimal near-feasible solutions. We refine the estimates of distances between epi-graphs in $\mathbb{R}^n$ provided by [24, Chapter 7] and [29], and also make them applicable to general metric spaces. Approximations of lsc functions on $\mathbb{R}^n$ by epi-splines is given by [27]. Here, we extend such approximations to lsc functions on separable metric spaces and proper metric spaces, and also give rates of convergence, which are novel even for $\mathbb{R}^n$. We refer to [10] for a general treatment of topologies on collections of closed sets; see also [1, 8, 24] for comprehensive descriptions of epi-convergence and its connections to variational analysis broadly.

Our motivation for going beyond normed linear spaces, which is the setting of [5, 6, 7], derives from emerging applications in nonparametric statistics, curve fitting, and stochastic processes that aim to identify an optimal function according to some criterion. A class of functions over which such optimal fitting might take place is the collection of lsc functions on $\mathbb{R}^n$, often simply with $n = 1$; see [30, 26, 27] for applications. The class of such lsc functions offers obvious modeling flexibility, which is important to practitioners, but under the aw-distance the class is a proper metric space that fails to be linear [24, Theorem 7.58]. Since it is proper, every closed ball in this metric space is compact and the existence of solutions of such optimal fitting problems is more easily established. We observe that the metric given to this class of lsc functions has the consequence that proximity of two functions implies closeness of their minimizers. This property is often important in probability density estimation, where the focus is on the modes of the density functions, i.e., the maximizers of the density functions. When fitting cumulative distribution functions, the metric metrizes weak convergence [29]. In both of these cases a reorientation towards upper semicontinuous functions instead of lsc functions is needed. In fact, nearly every result in this paper can be stated in terms of extended real-valued upper semincontinuous functions. However, we maintain the lsc perspective for simplicity.

There is an extensive literature on local stability of optimization problems under parametric perturbations; see for example [13, 16, 20, 19, 21, 18, 15] for a small collection of references. In contrast to these local stability results, dealing with "small" perturbations of an optimization problem, we present global results. That is, we give estimates of the distance between solutions of two problems that might be arbitrarily far apart. The ability to estimate the solution of one problem from that of another rather different problem is especially important in stochastic optimization, optimal control, and semi-infinite programming, and their numerous applications, as there we might only be able to construct and solve coarse approximations of the problem of interest.

The paper is organized as follows. After the review of epi-convergence in Section 2, we proceed in Section 3 with estimates of the aw-distance. Section 4 presents bounds on solution errors for optimization problems. Section 5 defines epi-splines and discusses their approximation properties.

2 Background

Throughout, we let (X, d) be a metric space and lsc-fcns$(X) := \{f : X \to \overline{I\!R} \;:\; f \text{ lsc and } f \not\equiv \infty\}$, where $\overline{I\!R} := I\!R \cup \{-\infty, \infty\}$. Thus, every $f \in$ lsc-fcns(X) has a nonempty closed epi-graph $\operatorname{epi} f := \{(x, x_0) \in X \times I\!R \;:\; f(x) \leq x_0\}$. For a nonempty closed $B \subseteq X$, we also write lsc-fcns(B) for the subset of lsc-fcns(X) consisting of functions f with $f(x) = \infty$ for all $x \notin B$. When considering $X \times I\!R$, we use the product metric $\bar{d}(\bar{x}, \bar{y}) := \max\{d(x, y), |x_0 - y_0|\}$ for $\bar{x} = (x, x_0) \in X \times I\!R$ and $\bar{y} = (y, y_0) \in X \times I\!R$. Let $I\!N := \{1, 2, ...\}$. Convergence is indicated by $\to$ regardless of type, with the being meaning clear from the context.

We recall (see for example [1, 8, 24]) that $f^\nu : X \to \overline{I\!R}$ *epi-converge* to $f : X \to \overline{I\!R}$ if and only if

$$\begin{aligned}
&\text{for every } x^\nu \to x, \operatorname{liminf} f^\nu(x^\nu) \geq f(x), \text{ and} \\
&\text{for some } x^\nu \to x, \operatorname{limsup} f^\nu(x^\nu) \leq f(x).
\end{aligned}$$

Epi-convergence neither implies nor is implied by pointwise convergence. Uniform convergence ensures epi-convergence, but fails to handle extended real-valued functions satisfactory—a necessity in constrained optimization problems.

For $f : X \to \overline{I\!R}$, $C \subset X$, and $\varepsilon \geq 0$, let $\inf f := \inf\{f(x) \;:\; x \in X\}$, $\inf_C f := \inf\{f(x) \;:\; x \in C\}$, $\operatorname{argmin} f := \{x \in X \;:\; f(x) = \inf f\}$, and $\varepsilon\text{-}\operatorname{argmin} f := \{x \in X \;:\; f(x) \leq \inf f + \varepsilon\}$. It is well known that epi-convergence ensures convergence of solutions of minimization problems (see for example [24, Chapter 7] and [3, 1]):

2.1 Proposition (convergence of minimizers) *Suppose that $f^\nu : X \to \overline{I\!R}$ epi-converges to $f : X \to \overline{I\!R}$. Then,*

$$\operatorname{limsup}\,(\inf f^\nu) \leq \inf f.$$

Moreover, if $x^k \in \operatorname{argmin} f^{\nu_k}$ and $x^k \to x$ for some increasing subsequence $\{\nu_1, \nu_2, ...\} \subset I\!N$, then $x \in \operatorname{argmin} f$ and $\lim_{k\to\infty} \inf f^{\nu_k} = \inf f$.

A strengthening of epi-convergence ensures the convergence of minima and approximation of minimizers (see for example [28]).

2.2 Definition (tight epi-convergence) *The functions $f^\nu : X \to \overline{I\!R}$ epi-converge tightly to $f : X \to \overline{I\!R}$ if they epi-converge to f and for all $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subseteq X$ and an integer ν_ε such that*

$$\inf_{K_\varepsilon} f^\nu \leq \inf f^\nu + \varepsilon \quad \text{for all } \nu \geq \nu_\varepsilon.$$

2.3 Proposition (convergence of infima) *Suppose that $f^\nu : X \to \overline{I\!R}$ epi-converges to $f : X \to \overline{I\!R}$ and $\inf f$ is finite. Then, f^ν epi-converges tightly to f*

(i) *if and only if* $\inf f^\nu \to \inf f$;

(ii) *if and only if there exists a sequence* $\varepsilon^\nu \searrow 0$ *such that* ε^ν-$\operatorname{argmin} f^\nu$ *set-converges*[1] *to* $\operatorname{argmin} f$.

Throughout, let $\tilde{x} \in X$ be fixed. The choice of $\tilde{x}$ can be made arbitrarily, but results might be sharper if $\tilde{x}$ is somewhat near minimizers of functions of interest as the analysis relies on the intersection of epi-graphs with $\mathbb{S}_\rho := I\!B_\rho \times [-\rho, \rho]$, where $I\!B_\rho := I\!B(\tilde{x}, \rho) := \{x \in X \ : \ d(\tilde{x}, x) \leq \rho\}$ and $\rho \geq 0$. The *Attouch-Wets (aw) distance* $d\!l$ (given $\tilde{x}$) is defined for any $f, g \in$ lsc-fcns(X) as

$$d\!l(f, g) := \int_0^\infty d\!l_\rho(f, g) e^{-\rho} d\rho,$$

where the *ρ-aw-distance*, $\rho \geq 0$, is given by

$$d\!l_\rho(f, g) := \sup\left\{ \left| \operatorname{dist}\left(\bar{x}, \operatorname{epi} f\right) - \operatorname{dist}\left(\bar{x}, \operatorname{epi} g\right) \right| \ : \ \bar{x} \in \mathbb{S}_\rho \right\},$$

with dist giving the usual point-to-set distance, which here is given by

$$\operatorname{dist}\left(\bar{x}, \bar{C}\right) = \inf\left\{ \bar{d}(\bar{x}, \bar{y}) \ : \ \bar{y} \in \bar{C} \right\} \text{ if } \bar{C} \subset X \times I\!R \text{ is nonempty and } \operatorname{dist}(\bar{x}, \emptyset) = \infty.$$

This setup resembles that of [24, Section 7.I], but there $X = I\!R^n$ and the Euclidean distance is used on $I\!R^n \times I\!R$. Generally, (lsc-fcns$(X), d\!l$) is a metric space [10, Section 3.1] that is complete whenever (X, d) is complete [10, Theorem 3.1.3]. We conclude from [10, Theorem 3.1.7] that for $f^\nu, f \in$ lsc-fcns(X),

$$d\!l(f^\nu, f) \to 0 \text{ implies that } f^\nu \text{ epi-converges to } f.$$

We recall that a metric space is *proper* if every closed ball in that space is compact. If (X, d) is proper, then the converse also holds: f^ν epi-converges to f implies that $d\!l(f^\nu, f) \to 0$ (see [10, Theorem 3.1.7] and [5, Theorem 4.2, Lemma 4.3], with the latter results being stated for $I\!R^n$ but their proofs carry over to the case of proper metric spaces). In addition to $X = I\!R^n$ with the usual metric, (X, d) is proper when $X =$ lsc-fcns$(I\!R^n)$ and $d = d\!l$. In fact, (lsc-fcns$(I\!R^n), d\!l$) is a proper complete separable metric space; [24, Theorem 7.58] and [27, Corollary 3.6]. This example is a motivation for the development due to applications in nonparametric statistics, curve fitting, and stochastic processes; see [26, 27, 30].

We use the following well-known fact repeatedly.

2.4 Lemma *If* (Y, d_Y) *is a proper metric space, then* $\operatorname{argmin}\{d_Y(y, y') \ : \ y' \in C\} \neq \emptyset$ *for every* $y \in Y$ *and nonempty closed set* $C \subset Y$.

Since the aw-distance quantifies epi-convergence, it is clear that its value for two lsc functions, or that of its estimates, leads to bounds on the distance between optimal solutions and optimal values of the two functions. Estimates of the aw-distance is the subject of the next section, with solutions being dealt with in Section 4.

[1] The outer limit of a sequence of sets $\{A^\nu\}_{\nu \in I\!N}$, denoted by $\limsup A^\nu$, is the collection of points x to which a subsequence of $\{x^\nu\}_{\nu \in I\!N}$, with $x^\nu \in A^\nu$, converges. The inner limit, denote by $\liminf A^\nu$, is the points to which a sequence of $\{x^\nu\}_{\nu \in I\!N}$, with $x^\nu \in A^\nu$, converges. If both limits exist and are identical to A, we say that A^ν set-converges to A; see [10, 24].

3 Distance Estimates

This section gives practically important estimates of the aw-distance between two lsc functions. We begin with defining an auxiliary quantity that estimates $d\!l_\rho$. For $\rho \geq 0$ and $f, g \in$ lsc-fcns(X), let

$$\hat{d\!l}_\rho(f,g) := \max\Big\{e\big(\operatorname{epi} f \cap \mathbb{S}_\rho, \operatorname{epi} g\big),\ e\big(\operatorname{epi} g \cap \mathbb{S}_\rho, \operatorname{epi} f\big)\Big\},$$

where the *excess* of a set C over a set D is given by

$$e(C, D) := \sup\{\operatorname{dist}(z, D)\ :\ z \in C\} \text{ if } C, D \text{ are nonempty},$$

$e(C, D) = \infty$ if C nonempty and D empty, and $e(C, D) = 0$ otherwise. Roughly speaking, $\hat{d\!l}_\rho(f,g)$ is the "padding" of epi g needed for it to contain epi $f \cap \mathbb{S}_\rho$ and vice versa. The relations among $d\!l$, $d\!l_\rho$, and $\hat{d\!l}_\rho$ given next extend [24, Exercise 7.60] from $X = I\!R^n$ to metric spaces.

3.1 Proposition (estimates of aw-distance) *For $f, g \in$ lsc-fcns(X), the following holds, where we use the notation $\delta_f = \operatorname{dist}((\tilde{x}, 0), \operatorname{epi} f)$ and similarly for g.*

(i) $d\!l_\rho(f,g)$ *and* $\hat{d\!l}_\rho(f,g)$ *are nondecreasing functions of* ρ;

(ii) $d\!l_{\rho'}(f,g) - d\!l_\rho(f,g) \leq 2e(\mathbb{S}_{\rho'}, \mathbb{S}_\rho)$ *for* $\rho' \geq \rho \geq 0$;

(iii) $\hat{d\!l}_\rho(f,g) \leq d\!l_\rho(f,g) \leq \hat{d\!l}_{\rho'}(f,g)$ *for* $\rho' > 2\rho + \max\{\delta_f, \delta_g\}$ *for* $\rho \geq 0$;

(iv) $d\!l_\rho(f,g) \leq \max\{\delta_f, \delta_g\} + \rho$ *for* $\rho \geq 0$;

(v) $d\!l(f,g) \geq (1 - e^{-\rho})|\delta_f - \delta_g| + e^{-\rho} d\!l_\rho(f,g)$ *for* $\rho \geq 0$;

(vi) $d\!l(f,g) \leq (1 - e^{-\rho}) d\!l_\rho(f,g) + e^{-\rho}[\max\{\delta_f, \delta_g\} + \rho + 1]$ *for* $\rho \geq 0$;

(vii) $|\delta_f - \delta_g| \leq d\!l(f,g) \leq \max\{\delta_f, \delta_g\} + 1$.

If (X, d) is proper, then $>$ can be replaced by $\geq$ in (iii).

Proof. See appendix. □

A nearly precise estimate of $\hat{d\!l}_\rho$ is provided by the following convenient quantity, which is closely related to the Kenmochi condition of [5]. For $f, g \in$ lsc-fcns(X), $\rho \geq 0$, and $\delta \geq 0$, let

$$\begin{aligned} \hat{d\!l}{}^{\delta}_\rho(f,g) := \inf\Big\{\eta \geq 0 :\ & \inf\nolimits_{B(x,\eta+\delta)} g \leq \max\{f(x), -\rho\} + \eta, \forall x \in \operatorname{lev}_\rho f \cap I\!B_\rho \\ & \inf\nolimits_{B(x,\eta+\delta)} f \leq \max\{g(x), -\rho\} + \eta, \forall x \in \operatorname{lev}_\rho g \cap I\!B_\rho\Big\}, \end{aligned}$$

where $\operatorname{lev}_\alpha f := \{x \in X\ :\ f(x) \leq \alpha\}$. Below, we also let $\operatorname{dom} f := \{x \in X\ :\ f(x) < \infty\}$. The next proposition extends a result in [29] from $X = I\!R^n$ to general metric spaces.

3.2 Proposition (estimates for auxiliary quantity) *For $f, g \in$ lsc-fcns(X), $\rho \in [0,\infty)$, and $\delta > 0$,*

$$\hat{d\!l}^{\delta}_{\rho}(f,g) \leq \hat{d\!l}_{\rho}(f,g) \leq \hat{d\!l}^{0}_{\rho}(f,g) < \infty.$$

If (X,d) is proper, then $\hat{d\!l}^{0}_{\rho}(f,g) = \hat{d\!l}_{\rho}(f,g)$.

Proof. See appendix. □

We state next an upper bound on $\hat{d\!l}^{0}_{\rho}$ in the case of Lipschitz continuous functions. We say that a function $f : X \to \overline{I\!R}$ is Lipschitz continuous with modulus $\kappa : [0,\infty) \to [0,\infty)$ (relative to $\tilde{x}$) if

$$|f(x) - f(y)| \leq \kappa(\rho) d(x,y) \text{ for all } x, y \in I\!B_{\rho} \text{ and } \rho \geq 0.$$

Parallel to $\hat{d\!l}_{\rho}(f,g)$, we also define with a slight abuse of notation

$$\hat{d\!l}_{\rho}(C,D) := \max\Big\{e\big(C \cap I\!B_{\rho}, D\big),\ e\big(D \cap I\!B_{\rho}, C\big)\Big\},\ \rho \geq 0,\ C, D \subset X \text{ nonempty closed.}$$

The next result generalizes a statement in [29] to metric spaces and also tightens it slightly. We define for any $C \subset X$ the function $\iota_C : X \to \overline{I\!R}$ that has $\iota_C(x) = 0$ if $x \in C$ and $\iota_C(x) = \infty$ otherwise. We also adopt the usual convention that $-\infty + \infty = \infty$.

3.3 Proposition (distance estimate for Lipschitz functions) *Suppose that $C, D \subset X$ are nonempty closed sets and $f, g : X \to I\!R$ are Lipschitz continuous with common modulus κ. Then, for $\rho \in [0,\infty)$ and $\rho' \in (\rho + \hat{d\!l}_{\rho}(C,D), \infty)$,*

$$\hat{d\!l}^{0}_{\rho}(f + \iota_C, g + \iota_D) \leq \sup\nolimits_{A_{\rho}} |f - g| + \max\{1, \kappa(\rho')\}\hat{d\!l}_{\rho}(C,D)$$

where $A_{\rho} = (\mathrm{lev}_{\rho}\{f + \iota_C\} \cup \mathrm{lev}_{\rho}\{g + \iota_D\}) \cap I\!B_{\rho}$.

If (X,d) is proper, then the result also holds for $\rho' = \rho + \hat{d\!l}_{\rho}(C,D)$.

Proof. See appendix. □

Sometimes the following asymmetric quantity is useful. For $f, g \in$ lsc-fcns(X) and $\rho \geq 0$, let

$$\eta^{+}_{\rho}(f;g) := \inf\big\{\eta \geq 0\ :\ \inf\nolimits_{B(x,\eta)} f \leq \max\{g(x), -\rho\} + \eta\ \ \forall x \in \mathrm{lev}_{\rho}\, g \cap I\!B_{\rho}\big\}.$$

Obviously, $\hat{d\!l}^{0}_{\rho}(f,g) = \max\{\eta^{+}_{\rho}(f;g), \eta^{+}_{\rho}(g;f)\}$ and $\eta^{+}_{\rho}(f;g) \leq \max\{0, \sup_{\mathrm{lev}_{\rho} g \cap I\!B_{\rho}} f - g\}$. Proposition 3.2 implies that for any $f, g \in$ lsc-fcns(X), $\eta^{+}_{\rho}(f;g) < \infty$.

4 Solution Estimates

Since $f, g \in$ lsc-fcns(X) completely define the problems $\min_{x \in X} f(x)$ and $\min_{x \in X} g(x)$, and $d\!l$ quantifies epi-convergence, it is clear that $d\!l(f,g)$ leads to estimates of $|\inf f - \inf g|$ as well as some notion of distance between $\mathrm{argmin}\, f$ and $\mathrm{argmin}\, g$. In this section, we provide such estimates as well as estimates between near-optimal solutions and near-optimal near-feasible solutions. Instead of $d\!l$, we work directly with the auxiliary quantity $\hat{d\!l}_\rho^0$, which is simpler to estimate in most practical situations; see for example Proposition 3.3. In view of Propositions 3.1 and 3.2, the difference between the two quantities are anyhow small for large ρ. We note that the results of this section are practically more useful when function values are scaled to be of the same order of magnitude as the parameter ρ.

We start by developing a result for optimal values that generalizes a statement in [6] by considering general metric spaces, permitting empty sets of optimal solutions, and dealing with the asymmetric quantity η_ρ^+.

4.1 Theorem (approximation of optimal value) *Suppose that $f, g \in$ lsc-fcns(X), $\rho \in (0, \infty)$, $\rho > \inf f \geq -\rho$, and ε-$\operatorname{argmin} f \cap I\!B_\rho \neq \emptyset$ for all $\varepsilon > 0$. Then,*

$$\inf g - \inf f \leq \eta_\rho^+(g; f) \leq \hat{d\!l}_\rho^0(f, g).$$

If the assumption about f also holds for g, then

$$|\inf g - \inf f| \leq \hat{d\!l}_\rho^0(f, g).$$

Proof. We note that $\eta_\rho^+(g; f) < \infty$. Let $\eta \in (\eta_\rho^+(g; f), \infty)$ be arbitrary. Then, for all $x \in \operatorname{lev}_\rho f \cap I\!B_\rho$,

$$\inf g \leq \inf_{I\!B(x,\eta)} g \leq \max\{f(x), -\rho\} + \eta. \tag{1}$$

Set $\varepsilon_0 = \rho - \inf f > 0$. Let $\varepsilon \in (0, \varepsilon_0]$ be arbitrary and $x_\varepsilon \in \varepsilon$-$\operatorname{argmin} f \cap I\!B_\rho$. Then, $f(x_\varepsilon) \leq \inf f + \varepsilon \leq \inf f + \varepsilon_0 = \rho$ and thus $x_\varepsilon \in \operatorname{lev}_\rho f$. Applying (1) with $x = x_\varepsilon$ results in

$$\inf g \leq \max\{f(x_\varepsilon), -\rho\} + \eta \leq \max\{\inf f + \varepsilon, -\rho\} + \eta \leq \inf f + \varepsilon + \eta.$$

After letting ε and η tend to their lower limits, we obtain that $\inf g \leq \inf f + \eta_\rho^+(g; f) \leq \inf f + \hat{d\!l}_\rho^0(f, g)$. The final result follows after a replication of these arguments with the roles of f and g reversed. ◻

We observe that if $\operatorname{argmin} f \cap I\!B_\rho \neq \emptyset$, then it suffices to have $\rho \geq \inf f \geq -\rho$ in Theorem 4.1.

To enable a statement about optimal solutions, we need to bring in conditioning. The next result, which generalizes a similar statement in [6] to metric spaces, carries this out.

4.2 Theorem (approximation of optimal solutions) *Suppose that $f, g \in$ lsc-fcns(X) are such that*

$$\inf f, \inf g \in [-\rho, \rho], \ \ \operatorname{argmin} f \cap I\!B_\rho \neq \emptyset, \ \ \text{and} \ \ \operatorname{argmin} g \cap I\!B_\rho \neq \emptyset$$

for some $\rho \in [0, \infty)$ and that there exists an increasing and continuous function $\psi_f : [0, \infty) \to [0, \infty)$, with $\psi_f(0) = 0$ such that

$$f(x) - \inf f \geq \psi_f\big(\operatorname{dist}(x, \operatorname{argmin} f)\big) \text{ for all } x \in X.$$

Then,

$$e\left(\operatorname{argmin} g \cap I\!B_\rho, \operatorname{argmin} f\right) \leq \hat{d\!l}_\rho^0(f, g) + \psi_f^{-1}\left(2\hat{d\!l}_\rho^0(f, g)\right).$$

Proof. From Proposition 3.2, $\hat{d\!l}^0_\rho(f,g) < \infty$. Let $\eta \in (\hat{d\!l}^0_\rho(f,g), \infty)$ be arbitrary. For all $x \in \operatorname{lev}_\rho g \cap I\!B_\rho$, $\inf_{\boldsymbol{B}(x,\eta)} f \le \max\{g(x), -\rho\} + \eta$. In view of the property of ψ_f and the fact that $\inf g \ge -\rho$, we find that for $x \in \operatorname{lev}_\rho g \cap I\!B_\rho$,

$$\begin{aligned} g(x) - \inf f + \eta = \max\{g(x), -\rho\} + \eta - \inf f &\ge \inf_{y \in \boldsymbol{B}(x,\eta)} f(y) - \inf f \\ &\ge \inf_{y \in \boldsymbol{B}(x,\eta)} \psi_f\big(\operatorname{dist}(y, \operatorname{argmin} f)\big). \end{aligned}$$

By Theorem 4.1, $\inf g - \inf f \le \eta$; see the comment after that theorem to establish that ρ is sufficently large despite the fact that it might coincide with $\inf f$ and $\inf g$. Thus, for $x \in \operatorname{argmin} g \cap I\!B_\rho$, which of course implies that $x \in \operatorname{lev}_\rho g$, we have that

$$\begin{aligned} 2\eta \ge \inf g - \inf f + \eta &= g(x) - \inf f + \eta \\ &\ge \inf_{y \in \boldsymbol{B}(x,\eta)} \psi_f\big(\operatorname{dist}(y, \operatorname{argmin} f)\big) \\ &\ge \psi_f\big(\inf_{y \in \boldsymbol{B}(x,\eta)} \operatorname{dist}(y, \operatorname{argmin} f)\big), \end{aligned}$$

where the last inequality follows from the increasing property of ψ_f. Therefore, we have that

$$\inf_{y \in \boldsymbol{B}(x,\eta)} \operatorname{dist}\big(y, \operatorname{argmin} f\big) \le \psi_f^{-1}(2\eta).$$

Let $\varepsilon > 0$ be arbitrary. There exists an $x_\varepsilon \in I\!B(x,\eta)$ such that

$$\operatorname{dist}\big(x_\varepsilon, \operatorname{argmin} f\big) \le \inf_{y \in \boldsymbol{B}(x,\eta)} \operatorname{dist}\big(y, \operatorname{argmin} f\big) + \varepsilon.$$

These facts then imply that for $x \in \operatorname{argmin} g \cap I\!B_\rho$,

$$\begin{aligned} \operatorname{dist}(x, \operatorname{argmin} f) \le \operatorname{dist}(x_\varepsilon, \operatorname{argmin} f) + d(x_\varepsilon, x) &\le \inf_{y \in \boldsymbol{B}(x,\eta)} \operatorname{dist}\big(y, \operatorname{argmin} f\big) + \varepsilon + \eta \\ &\le \psi_f^{-1}(2\eta) + \varepsilon + \eta. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we have that $\operatorname{dist}(x, \operatorname{argmin} f) \le \eta + \psi_f^{-1}(2\eta)$. This holds for all $x \in \operatorname{argmin} g \cap I\!B_\rho$ and consequently

$$e\big(\operatorname{argmin} g \cap I\!B_\rho, \operatorname{argmin} f\big) \le \eta + \psi_f^{-1}(2\eta).$$

Since ψ_f^{-1} is continuous, the conclusion follows by letting η tend to $\hat{d\!l}^0_\rho(f,g)$. □

The bound is sharpe even for $X = I\!R$ as demonstrated by the following example. Let $f(x) = x^2$ and for $\eta \ge 0$ let $g_\eta(x) = (x-\eta)^2$ if $x \in I\!R \setminus ([0, 2\eta] \cup \{\eta + \sqrt{2\eta}\})$, $g_\eta(x) = \eta$ if $x \in [0, 2\eta] \cup \{\eta + \sqrt{2\eta}\}$. One can show that for $\tilde{x} = 0$ and sufficiently large ρ, $\hat{d\!l}^0_\rho(f,g) = \eta$. Then, $\operatorname{argmin} f = \{0\}$ and $\operatorname{argmin} g_\eta = [0, 2\eta] \cup \{\eta + \sqrt{2\eta}\}$. Since the conditioning function $\psi_f(t) = t^2$ in this case, we see that the conclusion of the theorem holds with equality when $\eta \le 2$.

In practice, it is difficult to develop a conditioning function ψ_f as required by Theorem 4.2; see [6] for a thorough discussion. Fortunately, a strong Lipschitz-type statement can be made about near-optimal solutions without the knowledge about such conditioning. Even for $X = I\!R^n$, the next result is novel by considering different levels of near optimality for the two problems and avoiding the convexity assumption of [7] and [24, Theorem 7.69].

4.3 Theorem (approximation of near-optimal solutions) *Suppose that* $f, g \in$ *lsc-fcns*(X), $\rho \in (0, \infty)$, $\inf f \in [-\rho, \rho)$, *and* γ- $\operatorname{argmin} f \cap I\!B_\rho \neq \emptyset$ *for all* $\gamma > 0$. *If* $\varepsilon \geq 0$, $\delta > 0$, *and* $\inf g \in [-\rho, \rho - \varepsilon]$, *then*

$$e\big(\varepsilon\text{- argmin}\, g \cap I\!B_\rho, (\varepsilon + 2\bar{\eta} + \delta)\text{- argmin}\, f\big) \leq \bar{\eta}, \;\; \text{where } \bar{\eta} = \hat{d\!l}_\rho^0(f, g).$$

If in addition (X, d) *is proper, then* δ *can be set to zero.*

Proof. From Proposition 3.2, $\hat{d\!l}_\rho^0(f, g) < \infty$. Let $\eta \in (\hat{d\!l}_\rho^0(f, g), \infty)$ be arbitrary. For all $x \in \operatorname{lev}_\rho g \cap I\!B_\rho$, $\inf_{\boldsymbol{B}(x,\eta)} f \leq \max\{g(x), -\rho\} + \eta$. Let $x \in \varepsilon$- $\operatorname{argmin} g \cap I\!B_\rho$, which implies that $x \in \operatorname{lev}_\rho g$. Since $\inf g \geq -\rho$, we therefore have that

$$\inf_{\boldsymbol{B}(x,\eta)} f \leq \max\{g(x), -\rho\} + \eta \leq \max\{\inf g + \varepsilon, -\rho\} + \eta = \inf g + \varepsilon + \eta. \tag{2}$$

There exists $x_\delta \in I\!B(x, \eta)$ such that $f(x_\delta) \leq \inf_{\boldsymbol{B}(x,\eta)} f + \delta/2$. From above we then have that $f(x_\delta) \leq \inf g + \varepsilon + \eta + \delta/2 \leq \inf f + \varepsilon + 2\eta + \delta/2$, where the last inequality follows from Theorem 4.1. Thus, $x_\delta \in (\varepsilon + 2\eta + \delta/2)$- $\operatorname{argmin} f$ and $\operatorname{dist}(x, (\varepsilon + 2\eta + \delta/2)\text{- argmin}\, f) \leq \eta$. Since this holds for all $x \in \varepsilon$- $\operatorname{argmin} g \cap I\!B_\rho$,

$$e\big(\varepsilon\text{- argmin}\, g \cap I\!B_\rho, (\varepsilon + 2\eta + \delta/2)\text{- argmin}\, f\big) \leq \eta.$$

The first conclusion then follows after letting η tend to $\bar{\eta}$.

If (X, d) is proper, then we continue from (2) as follows. Since lsc functions attains their minimum over compact sets, there exists $\hat{x} \in \operatorname{argmin}_{\boldsymbol{B}(x,\eta)} f$ and thus $f(\hat{x}) \leq \inf g + \varepsilon + \eta \leq \inf f + \varepsilon + 2\eta$, where again the last inequality follows from Theorem 4.1. Consequently, $\hat{x} \in (\varepsilon + 2\eta)$- $\operatorname{argmin} f$ and $\operatorname{dist}(x, (\varepsilon + 2\eta)\text{- argmin}\, f) \leq \eta$. Let $\{\eta^\nu\}_{\nu \in I\!N}$ be such that $\eta^\nu \searrow \bar{\eta}$ and $\operatorname{dist}(x, (\varepsilon + 2\eta^\nu)\text{- argmin}\, f) \leq \eta^\nu$. In view of Lemma 2.4, there exists $y^\nu \in (\varepsilon + 2\eta^\nu)$- $\operatorname{argmin} f$ such that $d(x, y^\nu) = \operatorname{dist}(x, (\varepsilon + 2\eta^\nu)\text{- argmin}\, f)$. Since $\{y^\nu\}_{\nu \in I\!N}$ is contained in a ball, which under the additional assumption is compact, we have that there exists an $N \subset I\!N$ and $\bar{y}$ such that $y^\nu \to^N \bar{y}$. Since $d(x, y^\nu) \leq \eta^\nu$ for all ν, $d(x, \bar{y}) \leq \bar{\eta}$. Moreover, $f(y^\nu) \leq \varepsilon + 2\eta^\nu$ for all ν implies that $f(\bar{y}) \leq \varepsilon + 2\bar{\eta}$ because f is lsc. Thus, $\operatorname{dist}(x, (\varepsilon + 2\bar{\eta})\text{- argmin}\, f) \leq \bar{\eta}$. Since this holds for all $x \in \varepsilon$- $\operatorname{argmin} g \cap I\!B_\rho$, the second conclusion follows. □

The above bound is sharp even for $X = I\!R$. Suppose that $f, g : I\!R \to \overline{I\!R}$ are given by $f(x) = 1$ for $x \in [1, 2)$, $f(2) = -1$, and $f(x) = \infty$ otherwise, and $g(x) = 0$ for $x \in [0, 2]$ and $g(x) = \infty$ otherwise. Obviously, f, g are lsc. Let $\hat{d\!l}_\rho^0$ be defined with $\tilde{x} = 0$ and $\rho > 2$. Then, $\hat{d\!l}_\rho^0(f, g) = 1$. Clearly, $x = 0$ is in $\operatorname{argmin} g \cap I\!B_\rho$ and $\operatorname{dist}(x, 2\text{- argmin}\, f) = 1$. In fact, $e(\operatorname{argmin} g \cap I\!B_\rho, 2\text{- argmin}\, f) = 1$. Moreover, $\operatorname{dist}(x, \gamma\text{- argmin}\, f) = 2$ for $\gamma \in [0, 2)$.

Theorem 4.3 leads to the following corollary about rate of convergence.

4.4 Corollary (rate of convergence to near-optimal solutions) *Suppose that* $\rho \in (0, \infty)$, $\delta > \varepsilon \geq 0$, $f \in$ *lsc-fcns*(X), $\inf f \in [-\rho, \rho)$, *and* γ- $\operatorname{argmin} f \cap I\!B_\rho \neq \emptyset$ *for all* $\gamma > 0$. *If* $f^\nu \in$ *lsc-fcns*(X) *has* $\inf f^\nu \in [-\rho, \rho - \varepsilon]$ *for all* ν *and* $\hat{d\!l}_\rho^0(f^\nu, f) \to 0$, *then there exists an* $\bar{\nu}$ *such that*

$$e\big(\varepsilon\text{- argmin}\, f^\nu \cap I\!B_\rho, \delta\text{- argmin}\, f\big) \leq \hat{d\!l}_\rho^0(f^\nu, f) \;\; \text{for all } \nu \geq \bar{\nu}.$$

Proof. By Theorem 4.3 we have that for any $\delta_0 > 0$ and ν,

$$e\big(\varepsilon\text{-}\operatorname{argmin} f^\nu \cap I\!B_\rho, (\varepsilon + 2\eta^\nu + \delta_0)\text{-}\operatorname{argmin} f\big) \le \eta^\nu, \text{ where } \eta^\nu = \hat{d\!l}{}_\rho^0(f^\nu, f).$$

Set $\delta_0 = (\delta - \varepsilon)/2 > 0$. Since $\eta^\nu \to 0$, there exists an $\bar\nu$ such that $\eta^\nu \le (\delta - \varepsilon)/4$ for all $\nu \ge \bar\nu$. Since $\varepsilon + 2\eta^\nu + \delta_0 \le \varepsilon + 2(\delta - \varepsilon)/4 + (\delta - \varepsilon)/2 = \delta$ for such ν, we have that $e\big(\varepsilon\text{-}\operatorname{argmin} f^\nu \cap I\!B_\rho, \delta\text{-}\operatorname{argmin} f\big) \le \eta^\nu$ for $\nu \ge \bar\nu$, which establishes the conclusion. $\square$

Near-optimal solutions are feasible in the sense that $x \in \varepsilon\text{-}\operatorname{argmin} f$ and $\inf f < \infty$ implies that $x \in \operatorname{dom} f$. We also consider *near-feasibility*, which is often practically relevant, and reach the following results about level sets; see [7] for results about convex functions on normed linear spaces.

4.5 Theorem (approximation of level sets) *Suppose that $f, g \in$ lsc-fcns(X), $\rho \in [0, \infty)$, and $\delta \in [-\rho, \rho]$. Then, for any $\gamma > 0$,*

$$e(\operatorname{lev}_\delta g \cap I\!B_\rho, \operatorname{lev}_{\delta+\bar\eta+\gamma} f) \le \bar\eta = \eta_\rho^+(f; g)$$

If in addition (X, d) is proper, then γ can be set to zero.

Proof. We observe that $\eta_\rho^+(f; g) < \infty$. Let $\eta \in (\eta_\rho^+(f; g), \infty)$ be arbitrary. For all $x \in \operatorname{lev}_\rho g \cap I\!B_\rho$, $\inf_{I\!B(x,\eta)} f \le \max\{g(x), -\rho\} + \eta$. Let $x \in \operatorname{lev}_\delta g \cap I\!B_\rho$, which implies that $x \in \operatorname{lev}_\rho g$. Thus,

$$\inf_{I\!B(x,\eta)} f \le \max\{g(x), -\rho\} + \eta \le \max\{\delta, -\rho\} + \eta = \delta + \eta. \tag{3}$$

There exists $x_\gamma \in I\!B(x, \eta)$ such that $f(x_\gamma) \le \inf_{I\!B(x,\eta)} f + \gamma/2$. From above we then have that $f(x_\gamma) \le \delta + \eta + \gamma/2$. Consequently, $x_\gamma \in \operatorname{lev}_{\delta+\eta+\gamma/2} f$ and $\operatorname{dist}(x, \operatorname{lev}_{\delta+\eta+\gamma/2} f) \le \eta$. Since this holds for any $x \in \operatorname{lev}_\delta g \cap I\!B_\rho$, $e\big(\operatorname{lev}_\delta g \cap I\!B_\rho, \operatorname{lev}_{\delta+\eta+\gamma/2} f\big) \le \eta$. The first conclusion then follows after letting η tend to $\bar\eta$.

If (X, d) is proper, then we continue from (3) as follows. Since f attain its minimum over $I\!B(x, \eta)$ in this case, there exists $\hat x \in \operatorname{argmin}_{I\!B(x,\eta)} f$ and $f(\hat x) = \inf_{I\!B(x,\eta)} f \le \delta + \eta$. Thus, $\hat x \in \operatorname{lev}_{\delta+\eta} f$ and $\operatorname{dist}(x, \operatorname{lev}_{\delta+\eta} f) \le \eta$. Let $\{\eta^\nu\}_{\nu \in I\!N}$ be such that $\eta^\nu \searrow \bar\eta$ and $\operatorname{dist}(x, \operatorname{lev}_{\delta+\eta^\nu} f) \le \eta^\nu$. In view of Lemma 2.4, there exists $y^\nu \in \operatorname{lev}_{\delta+\eta^\nu} f$ such that $d(x, y^\nu) = \operatorname{dist}(x, \operatorname{lev}_{\delta+\eta^\nu} f)$. Since $\{y^\nu\}_{\nu \in I\!N}$ is contained in a ball, which under the additional assumption is compact, we have that there exists an $N \subset I\!N$ and $\bar y$ such that $y^\nu \to^N \bar y$. Since $d(x, y^\nu) \le \eta^\nu$ for all ν, $d(x, \bar y) \le \bar\eta$. Moreover, $f(y^\nu) \le \delta + \eta^\nu$ for all ν implies that $f(\bar y) \le \delta + \bar\eta$ because f is lsc. Thus, $\operatorname{dist}(x, \operatorname{lev}_{\delta+\bar\eta} f) \le \bar\eta$. Since this holds for any $x \in \operatorname{lev}_\delta g \cap I\!B_\rho$, the second conclusion follows. $\square$

When considering both near-optimality and near-feasibility, we adopt the following definition. For $\varepsilon, \delta \ge 0$, the set of near-optimal near-feasible solutions of the problem $\min\{f_0(x) \;:\; f(x) \le 0, x \in X\}$, which of course is equivalent to[2] $\min\{f_0 + \iota_{f \le 0}\}$, is given by

$$(\varepsilon, \delta)\text{-}\operatorname{argmin}\{f_0 + \iota_{f \le 0}\} := \{x \in X \;:\; f_0(x) \le \inf\{f_0 + \iota_{f \le 0}\} + \varepsilon, f(x) \le \delta\}.$$

The next results are the first ones dealing with near-optimality, near-feasibility, and the asymmetric quantity η_ρ^+ in a general setting.

[2]We use the slight abbreviation $\iota_{f \le 0}$ for $\iota_{\{x \in X : f(x) \le 0\}}$.

4.6 Theorem (approximation of near-optimal near-feasible solutions) *For* $\varepsilon, \rho, \eta, \eta_0 \in [0, \infty)$, *suppose that the functions* $f_0, g_0, f, g \in$ *lsc-fcns*(X) *satisfy*

(i) $\sup_{\boldsymbol{B}_\rho}\{g_0 - f_0\} \leq \eta_0$ *and* $\sup_{\boldsymbol{B}_\rho}\{g - f\} \leq \eta$

(ii) $\inf\{f_0 + \iota_{f \leq -\eta}\}$ *finite and* $\inf\{g_0 + \iota_{g \leq 0}\} \in [-\rho, \rho - \varepsilon]$

(iii) $\operatorname{argmin}\{f_0 + \iota_{f \leq -\eta}\} \cap \mathbb{B}_\rho \neq \emptyset$ *and* $\operatorname{argmin}\{g_0 + \iota_{g \leq 0}\} \cap \mathbb{B}_\rho \neq \emptyset$.

Then, for $\alpha, \delta, \geq 0$ *and* $\gamma > 0$,

$$e\big((\varepsilon, \delta)\text{-}\operatorname{argmin}\{g_0 + \iota_{g \leq 0}\} \cap \mathbb{B}_\rho, (\varepsilon + \eta^+ + \eta_0 + \gamma, \alpha)\text{-}\operatorname{argmin}\{f_0 + \iota_{f \leq -\eta}\}\big) \leq \eta^+,$$

where $\eta^+ = \eta^+_\rho(f_0 + \iota_{f \leq \alpha - \eta};\ g_0 + \iota_{g \leq \delta}) < \infty$.

If in addition (X, d) *is proper, then* γ *can be set to zero.*

Proof. First we note that $\eta^+_\rho(f_0 + \iota_{f \leq \alpha - \eta}; g_0 + \iota_{g \leq \delta})$ is finite because $f_0 + \iota_{f \leq \alpha - \eta}$ and $g_0 + \iota_{g \leq \delta}$ are in lsc-fcns(X). Second, let $\bar{\eta} \in (\eta^+, \infty)$. Then,

$$\inf_{\boldsymbol{B}(x, \bar{\eta})}\{f_0 + \iota_{f \leq \alpha - \eta}\} \leq \max\{g_0(x) + \iota_{g \leq \delta}(x), -\rho\} + \bar{\eta} \text{ for all } x \in \operatorname{lev}_\rho\{g_0 + \iota_{g \leq \delta}\} \cap \mathbb{B}_\rho.$$

For $x \in \mathbb{B}_\rho$ satisfying $g_0(x) \leq \inf\{g_0 + \iota_{g \leq 0}\} + \varepsilon$ and $g(x) \leq \delta$, we have $x \in \operatorname{lev}_\rho\{g_0 + \iota_{g \leq \delta}\}$. Since $\inf\{g_0 + \iota_{g \leq 0}\} \geq -\rho$,

$$\begin{aligned}\inf_{\boldsymbol{B}(x, \bar{\eta})}\{f_0 + \iota_{f \leq \alpha - \eta}\} &\leq \max\{g_0(x) + \iota_{g \leq \delta}(x), -\rho\} + \bar{\eta} \leq \max\{g_0(x), -\rho\} + \bar{\eta}\\ &\leq \inf\{g_0 + \iota_{g \leq 0}\} + \varepsilon + \bar{\eta} = \inf_{\boldsymbol{B}_\rho}\{g_0 + \iota_{g \leq 0}\} + \varepsilon + \bar{\eta} < \infty,\end{aligned} \tag{4}$$

where the equality follows from the fact that $\operatorname{argmin}\{g_0 + \iota_{g \leq 0}\} \cap \mathbb{B}_\rho \neq \emptyset$. We next consider two cases. First, if $\inf_{\boldsymbol{B}(x, \bar{\eta})}\{f_0 + \iota_{f \leq \alpha - \eta}\}$ is finite, then there exists $y \in \mathbb{B}(x, \bar{\eta})$ such that

$$f_0(y) + \iota_{f \leq \alpha - \eta}(y) \leq \inf_{\boldsymbol{B}(x, \bar{\eta})}\{f_0 + \iota_{f \leq \alpha - \eta}\} + \gamma/2.$$

Consequently,

$$\begin{aligned}f_0(y) + \iota_{f \leq \alpha - \eta}(y) &\leq \inf_{\boldsymbol{B}_\rho}\{g_0 + \iota_{g \leq 0}\} + \varepsilon + \bar{\eta} + \gamma/2\\ &\leq \inf_{\boldsymbol{B}_\rho}\{f_0 + \iota_{g \leq 0}\} + \varepsilon + \bar{\eta} + \eta_0 + \gamma/2\\ &\leq \inf_{\boldsymbol{B}_\rho}\{f_0 + \iota_{f \leq -\eta}\} + \varepsilon + \bar{\eta} + \eta_0 + \gamma/2,\end{aligned}$$

where the last inequality follows from the fact that $f(z) \leq -\eta$ implies that $g(z) \leq 0$ for $z \in \mathbb{B}_\rho$. Since the right-hand side above is finite, $f(y) \leq \alpha - \eta$ and $f_0(y) \leq \inf_{\boldsymbol{B}_\rho}\{f_0 + \iota_{f \leq -\eta}\} + \varepsilon + \bar{\eta} + \eta_0 + \gamma/2$. Since $\inf_{\boldsymbol{B}_\rho}\{f_0 + \iota_{f \leq -\eta}\} = \inf\{f_0 + \iota_{f \leq -\eta}\}$, we then have that

$$y \in (\varepsilon + \bar{\eta} + \eta_0 + \gamma/2, \alpha)\text{-}\operatorname{argmin}\{f_0 + \iota_{f \leq -\eta}\}.$$

Hence,

$$\operatorname{dist}\big(x, (\varepsilon + \bar{\eta} + \eta_0 + \gamma/2, \alpha)\text{-}\operatorname{argmin}\{f_0 + \iota_{f \leq -\eta}\}\big) \leq \bar{\eta}. \tag{5}$$

Second, we consider the other case when $\inf_{\mathbb{B}(x,\bar\eta)}\{f_0 + \iota_{f\leq\alpha-\eta}\} = -\infty$. Then, there exists $y \in \mathbb{B}(x,\bar\eta)$ such that again $f(y) \leq \alpha - \eta$ and $f_0(y) \leq \inf_{\mathbb{B}_\rho}\{f_0 + \iota_{f\leq-\eta}\} + \varepsilon + \bar\eta + \eta_0 + \gamma/2$. Thus, also in this case, (5) holds. Since this argument is valid for all $x \in \mathbb{B}_\rho$ satisfying $g_0(x) \leq \inf\{g_0 + \iota_{g\leq 0}\} + \varepsilon$ and $g(x) \leq \delta$,

$$e\big((\varepsilon,\delta)\text{- argmin}\{g_0 + \iota_{g\leq 0}\} \cap \mathbb{B}_\rho, (\varepsilon + \bar\eta + \eta_0 + \gamma/2, \alpha)\text{- argmin}\{f_0 + \iota_{f\leq-\eta}\}\big) \leq \bar\eta.$$

The main conclusion then follows after letting $\bar\eta$ tend to its lower limit of η^+.

If (X,d) is proper, then we continue from (4) by recognizing that there exists $y \in \operatorname{argmin}_{\mathbb{B}(x,\bar\eta)}\{f_0 + \iota_{f\leq\alpha-\eta}\}$ and thus

$$f_0(y) + \iota_{f\leq\alpha-\eta}(y) \leq \inf_{\mathbb{B}_\rho}\{g_0 + \iota_{g\leq 0}\} + \varepsilon + \bar\eta \leq \inf_{\mathbb{B}_\rho}\{f_0 + \iota_{f\leq-\eta}\} + \bar\eta + \eta_0 + \varepsilon < \infty.$$

Consequently, $f(y) \leq \alpha - \eta$ and $f_0(y) \leq \inf_{\mathbb{B}_\rho}\{f_0 + \iota_{f\leq-\eta}\} + \bar\eta + \eta_0 + \varepsilon$, or equivalently, $y \in (\varepsilon + \bar\eta + \eta_0, \alpha)\text{- argmin}\{f_0 + \iota_{f\leq-\eta}\}$. Following the same argument as used towards the end of the proof of Theorem 4.3, we find that the relation also holds as $\bar\eta$ reaches its lower limit of η^+ and

$$\operatorname{dist}\big(x, (\varepsilon + \eta^+ + \eta_0, \alpha)\text{- argmin}\{f_0 + \iota_{f\leq-\eta}\}\big) \leq \eta^+.$$

Since this holds for all $x \in \mathbb{B}_\rho$ with $g_0(x) \leq \inf\{g_0 + \iota_{g\leq 0}\} + \varepsilon$ and $g(x) \leq \delta$, the conclusion follows.$\square$

We next give an estimate of η_ρ^+ analogous to Proposition 3.3.

4.7 Proposition (bounds for Lipschitz continuous objective) *Suppose that $f, f_0, g, g_0 \in$ lsc-fcns(X), with f_0 being Lipschitz continuous functions with modulus κ, and $\inf\{f_0 + \iota_{f\leq 0}\}$ is finite. Then, for $\alpha, \delta, \rho \in [0,\infty)$ and $\rho' \in (\rho + e(\operatorname{lev}_\delta g \cap \mathbb{B}_\rho, \operatorname{lev}_\alpha f), \infty)$.*

$$\eta_\rho^+(f_0 + \iota_{f\leq\alpha}; g_0 + \iota_{g\leq\delta}) \leq \max\{0, \sup_{A_\rho^0} f_0 - g_0\} + \max\{1, \kappa(\rho')\} e(\operatorname{lev}_\delta g \cap \mathbb{B}_\rho, \operatorname{lev}_\alpha f),$$

where $A_\rho^0 = \operatorname{lev}_\rho\{g_0 + \iota_{g\leq\delta}\} \cap \mathbb{B}_\rho$.

Moreover, if in addition $\delta \leq \rho$, and $\alpha > \eta_\rho^+(f;g) + \delta$, then

$$e(\operatorname{lev}_\delta g \cap \mathbb{B}_\rho, \operatorname{lev}_\alpha f) \leq \eta_\rho^+(f;g).$$

If (X,d) is proper, $\alpha = \eta_\rho^+(f;g) + \delta$ is permitted.

Proof. We observe that $\eta_\rho^+(f_0 + \iota_{f\leq\alpha}; g_0 + \iota_{g\leq\delta})$ is finite even if $g_0 + \iota_{g\leq\delta}$ is identical to ∞ and that $\operatorname{lev}_\alpha f \neq \emptyset$. Thus, $e(\operatorname{lev}_\delta f \cap \mathbb{B}_\rho, \operatorname{lev}_\alpha f) < \infty$. Let $\rho' \in (\rho + e(\operatorname{lev}_\delta g \cap \mathbb{B}_\rho, \operatorname{lev}_\alpha f), \infty)$, and $\varepsilon \in (0, \rho' - \rho - e(\operatorname{lev}_\delta g \cap \mathbb{B}_\rho, \operatorname{lev}_\alpha f)]$. Set

$$\eta_\varepsilon = \max\{0, \sup_{A_\rho^0} f_0 - g_0\} + \max\{1, \kappa(\rho')\}\big[e(\operatorname{lev}_\delta g \cap \mathbb{B}_\rho, \operatorname{lev}_\alpha f) + \varepsilon\big].$$

Suppose that $x \in \operatorname{lev}_\rho\{g_0 + \iota_{g\leq\delta}\} \cap \mathbb{B}_\rho$, which implies that $g(x) \leq \delta$. Since $\operatorname{lev}_\alpha f \neq \emptyset$, there exists $y \in X$ such that $f(y) \leq \alpha$ and $d(x,y) \leq \inf\{d(x,y') \;:\; f(y') \leq \alpha, y' \in X\} + \varepsilon$. Thus,

$$e(\operatorname{lev}_\delta g \cap \mathbb{B}_\rho, \operatorname{lev}_\alpha f) \geq \operatorname{dist}(x, \operatorname{lev}_\alpha f) \geq d(x,y) - \varepsilon.$$

Consequently, $d(x,y) \le \eta_\varepsilon$ and $d(\tilde{x},y) \le d(\tilde{x},x)+d(x,y) \le \rho+e(\mathrm{lev}_\delta\, g\cap \mathbb{B}_\rho, \mathrm{lev}_\alpha\, f)+\varepsilon \le \rho'$. Combining these facts, we find that

$$\begin{aligned}\inf\nolimits_{\boldsymbol{B}(x,\eta_\varepsilon)}\{f_0+\iota_{f\le\alpha}\} &\le f_0(y) = f_0(y)-f_0(x)+f_0(x)-g_0(x)+g_0(x)\\ &\le \kappa(\rho')d(x,y)+\max\{0,\sup\nolimits_{A_\rho^0} f_0-g_0\}+\max\{g_0(x)+\iota_{g\le\delta}(x),-\rho\}\\ &\le \eta_\varepsilon+\max\{g_0(x)+\iota_{g\le\delta}(x),-\rho\}.\end{aligned}$$

Thus, $\eta_\rho^+(f_0+\iota_{f\le\alpha}; g_0+\iota_{g\le\delta}) \le \eta_\varepsilon$. Since this fact holds for all $\varepsilon>0$, the first conclusion follows.

An application of Theorem 4.5 results in the second part of the result. □

We obtain a rate of convergence result by combining Theorems 4.5, 4.6, and Proposition 4.7.

4.8 Proposition (rate of convergence to near-optimal near-feasible solutions) *For $\beta>\varepsilon\ge 0$, $\alpha>\delta\ge 0$, $\rho'>\rho\ge\delta$, and $\gamma\in(0,\alpha-\delta)$, suppose that the functions $f_0, f_0^\nu, f, f^\nu \in$ lsc-fcns(X), with f_0 being Lipschitz continuous with modulus κ, satisfy*

(i) $\inf\{f_0+\iota_{f\le 0}\}$, $\inf\{f_0+\iota_{f\le-\gamma}\}$ *are finite*

(ii) $\mathrm{argmin}\{f_0+\iota_{f\le-\gamma}\}\cap \mathbb{B}_\rho \neq \emptyset$

(iii) $\inf\{f_0^\nu+\iota_{f^\nu\le 0}\}\in[-\rho,\rho-\varepsilon]$ *for all* ν

(iv) $\mathrm{argmin}\{f_0^\nu+\iota_{f^\nu\le 0}\}\cap \mathbb{B}_\rho \neq \emptyset$ *for all* ν.

If $\sup_{\boldsymbol{B}_\rho}|f_0^\nu-f_0|\to 0$ *and* $\sup_{\boldsymbol{B}_\rho}|f^\nu-f|\to 0$, *as* $\nu\to\infty$, *then there exists an* $\bar\nu$ *such that for all* $\nu\ge\bar\nu$,

$$\begin{aligned}&e\big((\varepsilon,\delta)\text{-}\mathrm{argmin}\{f_0^\nu+\iota_{f^\nu\le 0}\}\cap\mathbb{B}_\rho, (\beta,\alpha)\text{-}\mathrm{argmin}\{f_0+\iota_{f\le-\gamma}\}\big)\\ &\le \sup\nolimits_{\boldsymbol{B}_\rho}|f_0-f_0^\nu|+\max\{1,\kappa(\rho')\}\sup\nolimits_{\boldsymbol{B}_\rho}|f^\nu-f|.\end{aligned}$$

Proof. There exists an $\bar\nu$ such that for all $\nu\ge\bar\nu$, $\sup_{\boldsymbol{B}_\rho}|f_0^\nu-f_0|\le(\beta-\varepsilon)/4$ and

$$\sup\nolimits_{\boldsymbol{B}_\rho}|f^\nu-f| \le \min\left\{\gamma, \frac{\rho'-\rho}{2}, \frac{\alpha-\delta-\gamma}{2}, \frac{\beta-\varepsilon}{4\max\{1,\kappa(\rho')\}}\right\}.$$

Let $\nu\ge\bar\nu$. We start with an application of Theorem 4.6 with f_0^ν and f^ν in the role of g_0 and g, respectively, and $\eta=\gamma$ and that theorem's γ being set to $(\beta-\varepsilon)/4$. This results in

$$e\big((\varepsilon,\delta)\text{-}\mathrm{argmin}\{f_0^\nu+\iota_{f^\nu\le 0}\}\cap\mathbb{B}_\rho, (\varepsilon+\eta^++(\beta-\varepsilon)/4+(\beta-\varepsilon)/4,\alpha)\text{-}\mathrm{argmin}\{f_0+\iota_{f\le-\gamma}\}\big)\le\eta^+, \quad (6)$$

where $\eta^+=\eta_\rho^+(f_0+\iota_{f\le\alpha-\gamma};\ f_0^\nu+\iota_{f^\nu\le\delta})$. Next, we invoke Theorem 4.5 and conclude that

$$e(\mathrm{lev}_\delta\, f^\nu\cap\mathbb{B}_\rho, \mathrm{lev}_{\alpha-\gamma}\, f)\le\eta_\rho^+(f;f^\nu)\le\sup\nolimits_{\boldsymbol{B}_\rho}|f^\nu-f|\le\frac{\rho'-\rho}{2}$$

because $\alpha - \gamma > \delta + (\alpha - \delta - \gamma)/2 \geq \delta + \eta^+_\rho(f; f^\nu)$. Finally, we bring in Proposition 4.7. Since $\rho + e(\mathrm{lev}_\delta f^\nu \cap I\!B_\rho, \mathrm{lev}_{\alpha-\gamma} f) \leq \rho + (\rho' - \rho)/2 < \rho'$, we conclude that the present choice of ρ' suffices in that proposition and α there is set to $\alpha - \gamma > 0$. Thus,

$$
\begin{aligned}
\eta^+ &\leq \sup\nolimits_{\boldsymbol{B}_\rho} |f_0^\nu - f_0| + \max\{1, \kappa(\rho')\} e(\mathrm{lev}_\delta f^\nu \cap I\!B_\rho, \mathrm{lev}_{\alpha-\gamma} f) \\
&\leq \sup\nolimits_{\boldsymbol{B}_\rho} |f_0^\nu - f_0| + \max\{1, \kappa(\rho')\} \sup\nolimits_{\boldsymbol{B}_\rho} |f^\nu - f| \\
&\leq \frac{\beta - \varepsilon}{4} + \max\{1, \kappa(\rho')\} \frac{\beta - \varepsilon}{4 \max\{1, \kappa(\rho')\}} \leq \frac{\beta - \varepsilon}{2}.
\end{aligned}
\tag{7}
$$

Thus, $\varepsilon + \eta^+ + (\eta - \varepsilon)/4 + (\eta - \varepsilon)/4 \leq \beta$ and we see from (6) and (7) that the conclusion holds. $\square$

We note that the proposition makes a statement about rate of convergence of near-optimal near-feasible solutions of the approximate problem $\min\{f_0^\nu + \iota_{f^\nu \leq 0}\}$ to solutions of a slightly restricted "original" problem $\min\{f_0 + \iota_{f \leq -\gamma}\}$, with $\gamma > 0$ arbitrarily small. The use of such a restriction allows us to avoid possibly hard-to-verify conditions on the constraint function and its level sets.

5 Epi-Splines and Construction of Approximations

In the previous sections, we bounded the aw-distance between two given lsc functions and related such bounds to solution estimates for the minimization problems defined by those functions. We now turn to the *construction* of a function that approximates a given lsc function. In practice, approximations of optimization problems depend on the nature of the application. We take an abstract perspective and examine piecewise constant functions that resemble the simple functions of integration theory and (zeroth-degree) polynomial splines from functional approximation theory. The approximating functions are defined by a finite number of parameters. As we see below, they can be made to approximate to an arbitrary level of accuracy any functions in lsc-fcns(X) under some assumptions on X, relying on epi-convergence and the aw-distance to formalize the meaning of accuracy. The results in this section certainly open up computational possibilities for solving difficulty optimization problems, but also provide new means to establish theoretical results about lsc functions through their finite-dimensional approximations.

We adopt the notation $\mathrm{cl}\, A$ and $\mathrm{int}\, A$ for the closure and interior, respectively, of a subset A of a topological space. The approximating functions are defined in terms a *finite* collection of subsets of X.

5.1 Definition (partition) *A finite collection $R_1, R_2, ..., R_N$ of open subsets of X is a partition of a closed set $B \subseteq X$ if $\cup_{k=1}^N \mathrm{cl}\, R_k = B$ and $R_k \cap R_l = \emptyset$ for all $k \neq l$.*

For any $f : X \to \overline{I\!R}$ and $x \in X$, let $\mathrm{liminf}_{x' \to x} f(x') := \lim_{\delta \downarrow 0} \inf_{x' \in \boldsymbol{B}(x,\delta)} f(x')$. Clearly, f is lsc if $\mathrm{liminf}_{x' \to x} f(x') \geq f(x)$ for all $x \in X$ (see for example [3, Section 2]). The approximating functions, called epi-splines, are defined next.

5.2 Definition (epi-splines) *An epi-spline $s : X \to \overline{I\!R}$, with partition $\mathcal{R} = \{R_k\}_{k=1}^N$ of a closed set*

$B \subseteq X$, *is a function that*

$$\text{on each } R_k,\ k = 1, ..., N, \text{ is constant,}$$
$$\text{has } s(x) = \infty \text{ for } x \notin B, \text{ and for every } x \in X, \text{ has } s(x) = \text{liminf}_{x' \to x}\, s(x').$$

The family of all such epi-splines is denoted by e-spl($\mathcal{R}$).

This definition straightforwardly extends from $I\!\!R^n$ to general metric spaces the one in [27]. There we also deal with "higher-order" epi-splines involving polynomials of degrees greater than zero on each R_k, which motivates the reference to "splines" in the name. The same possibility exists here, but we shy away from that subject due to the complications related to extending polynomials to general metric spaces. The reference to "epi" in the name is motivated by the choice of epi-convergence as the notion of convergence as we see below.

Clearly, by definition, every epi-spline is lsc. The ability of epi-splines to approximate arbitrary lsc functions relies on a refinement of the partition.

5.3 Definition (infinite refinement) *A sequence* $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ *of partitions of a closed set* $B \subseteq X$, *with* $\mathcal{R}^\nu = \{R_k^\nu\}_{k=1}^{N^\nu}$, *is an infinite refinement of* B *if*

$$\text{for every } x \in B \text{ and } \varepsilon > 0, \text{ there exist } \bar{\nu} \in I\!\!N \text{ and } \delta \in (0, \varepsilon) \text{ such that}$$
$$R_k^\nu \subset I\!\!B(x, \varepsilon) \text{ for every } \nu \geq \bar{\nu} \text{ and } k \text{ satisfying } R_k^\nu \cap I\!\!B(x, \delta) \neq \emptyset.$$

We note that this notion of refinement is local in nature, which is essential as we aim to address partitions of unbounded sets. A sufficient condition for the existence of an infinite refinement is separability.

5.4 Proposition (existence of infinite refinement) *If* $B \subseteq X$ *is nonempty, separable, and solid*[3], *then there exists an infinite refinement of* B.

Proof. Let $x^0 \in B$ and $B^\nu = I\!\!B(x^0, \nu) \cap B$, $\nu \in I\!\!N$. The separability of B implies that there exists a sequence $\{I\!\!B(x^j, \varepsilon^j)\}_{j \in I\!\!N}$, with $\{x^j\}_{j \in I\!\!N}$ a dense subset of B and $\{\varepsilon^j\}_{j \in I\!\!N}$ a dense subset of $(0, \infty)$ such that $\{(x^j, \varepsilon^j)\}_{j \in I\!\!N}$ is dense in $B \times (0, \infty)$ under the product topology. For every ν, the boundedness of B^ν implies that there exists a $J^\nu < \infty$ such that $B^\nu \subset \{I\!\!B(x^j, \varepsilon^j)\}_{j=1}^{J^\nu}$. Let $\{M^\nu\}_{\nu \in I\!\!N}$ be a sequence of scalars that tend to infinity and $M^\nu \geq J^\nu$.

We are then ready to construct the open sets that form the partitions, which subsequently will be shown to be an infinite refinement. For every ν the process is identical. First, sort the balls $\{I\!\!B(x^j, \varepsilon^j)\}_{j=1}^{M^\nu}$ in the order of nondecreasing radii and let this ordered set be $\{I\!\!B(x^{j_k^\nu}, \varepsilon^{j_k^\nu})\}_{k=1}^{M^\nu}$, i.e., $\varepsilon^{j_k^\nu} \leq \varepsilon^{j_{k+1}^\nu}$ for all $k = 1, ..., M^\nu - 1$. Second, set $R_1^\nu = \text{int}(B^\nu \cap I\!\!B(x^{j_1^\nu}, \varepsilon^{j_1^\nu}))$ and recursively

$$R_k^\nu = \text{int}\left(\left(B^\nu \cap I\!\!B(x^{j_k^\nu}, \varepsilon^{j_k^\nu})\right) \setminus \cup_{l=1}^{k-1} R_l^\nu\right) \text{ for } k = 2, 3, ..., M^\nu.$$

Set $N^\nu = M^\nu + 1$ and let $R_{N^\nu}^\nu = \text{int}(B \setminus I\!\!B(x^0, \nu))$. We observe that some R_k^ν maybe empty, but that is immaterial. Obviously, R_k^ν, $k = 1, ..., N^\nu$, are open and nonoverlapping, and $\cup_{k=1}^{N^\nu} R_k^\nu \subset B$. Since B

[3] B is solid if $B = \text{cl}(\text{int}\, B)$

is closed, we also have that $\cup_{k=1}^{N^\nu} \mathrm{cl}\, R_k^\nu \subset B$. We can conclude that R_k^ν, $k = 1, ..., N^\nu$, is a partition of B after establishing that $\cup_{k=1}^{N^\nu} \mathrm{cl}\, R_k^\nu \supset B$. Suppose that $x \in B$. Since $B = \mathrm{cl}(\mathrm{int}\, B)$ (i.e., is solid), there exist $\{y^\mu\}_{\mu \in I\!N}$ and positive numbers $\{\delta^\mu\}_{\mu \in I\!N}$ such that $y^\mu \to x$, $\delta^\mu \searrow 0$, and $I\!B(y^\mu, \delta^\mu) \subset \mathrm{int}\, B$. For every μ, there is a $k_\mu \in \{1, ..., N^\nu\}$ such that $I\!B(y^\mu, \delta^\mu) \cap R_{k_\mu}^\nu \neq \emptyset$ as we see from the construction. Hence, there exists $z^\mu \in I\!B(y^\mu, \delta^\mu) \cap R_{k_\mu}^\nu$, which, due to $\delta^\mu \searrow 0$, tend to x. This implies that $x \in \cup_{k=1}^{N^\nu} \mathrm{cl}\, R_k^\nu$ and therefore $\{R_k^\nu\}_{k=1}^{N^\nu}$ is a partition of B. This holds for all ν.

We next show that Definition 5.3 holds. Let $x \in B$ and $\varepsilon > 0$. There exists a $\bar\nu \geq d(x^0, x) + \varepsilon$ such that the collection $\{I\!B(x^j, \varepsilon^j)\}_{j=1}^{M^{\bar\nu}}$ contains a ball $I\!B(x^{j_*}, \varepsilon^{j_*}) \supset I\!B(x, \varepsilon/4)$ and $\varepsilon^{j_*} \leq \varepsilon/3$. Suppose that this ball is number k_*^ν in the sorted collections $\{I\!B(x^{j_k^\nu}, \varepsilon^{j_k^\nu})\}_{k=1}^{M^\nu}$, $\nu = \bar\nu, \bar\nu + 1, ...$ Consequently, $\cup_{k=1}^{k_*^\nu} \mathrm{cl}\, R_k^\nu \supset I\!B(x, \varepsilon/4)$ for all $\nu \geq \bar\nu$. By construction $R_k^\nu \subset I\!B(x^{j_k^\nu}, \varepsilon^{j_k^\nu})$ for all k and ν. Thus, for every $\nu \geq \bar\nu$ and $k = 1, ..., k_*^\nu$, $\sup_{y, y' \in R_k^\nu} d(y, y') \leq \varepsilon/3$ due to the nondecreasing radii of the balls in the sorted collections. We therefore have that $R_k^\nu \subset I\!B(x, \varepsilon)$ for every $k = 1, ..., k_*^\nu$ and $\nu \geq \bar\nu$ satisfying $R_k^\nu \cap I\!B(x, \varepsilon/4) \neq \emptyset$. Since from above we know that $\cup_{k=1}^{k_*^\nu} \mathrm{cl}\, R_k^\nu \supset I\!B(x, \varepsilon/4)$ for all $\nu \geq \bar\nu$, we conclude that $R_k^\nu \cap I\!B(x, \varepsilon/4) = \emptyset$ for all $k = k_*^\nu + 1, k_*^\nu + 2, ..., N^\nu$ and $\nu \geq \bar\nu$. Consequently, Definition 5.3 holds with $\delta = \varepsilon/4$. □

The above proof provides guidance towards the construction of infinite refinements, for which there are, of course, many possibilities. A main approximation results for epi-splines is given next, where the approximation is in the sense of epi-convergence and pointwise convergence. The result is an improvement over one in [27] by allowing X to be a general metric space, not only $I\!R^n$, and by also establishing pointwise convergence as well as an upper bound. Later, we give a stronger conclusion of convergence in the sense of the aw-distance under additional assumptions.

5.5 Theorem (approximation of lsc functions) *If $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ is an infinite refinement of a nonempty closed set $B \subseteq X$, then for every $f \in$ lsc-fcns(B) there exist epi-splines $s^\nu \in$ e-spl$(\mathcal{R}^\nu)$ satisfying the following:*

(i) s^ν epi-converges to f,

(ii) s^ν converges pointwise to f on X, and

(iii) $s^\nu(x) \leq \max\{-\nu, f(x)\}$ for all ν and $x \in X$.

Proof. Let $f \in$ lsc-fcns(B) and $\mathcal{R}^\nu = \{R_k^\nu\}_{k=1}^{N^\nu}$. For every $\nu \in I\!N$ and $k \in \{1, 2, ..., N^\nu\}$, we let

$$\sigma_k^\nu = \min\left\{\nu, \max\left\{-\nu, \inf_{\mathrm{cl}\, R_k^\nu} f\right\}\right\}$$

and construct $s^\nu : X \to \overline{I\!R}$ as follows:

$$s^\nu(x) = \min_{k=1,...,N^\nu} \left\{\sigma_k^\nu \ : \ x \in \mathrm{cl}\, R_k^\nu\right\} \text{ for } x \in B, \text{ and } s^\nu(x) = \infty \text{ for } x \notin B.$$

For every $\nu \in I\!N$, the open sets R_k^ν, $k = 1, ..., N^\nu$, are disjoint. Thus, if $x, x' \in R_k^\nu$, then $s^\nu(x) = s^\nu(x') = \sigma_k^\nu$, and s^ν is constant on R_k^ν. If $x \in B$, but not in R_k^ν for any $k = 1, ..., N^\nu$, then for the set

$K^\nu(x) = \{k \ : \ x \in \operatorname{cl} R_k^\nu\} \subset \{1, ..., N^\nu\}$ we have that $s^\nu(x) = \min\{\sigma_k^\nu \ : \ k \in K^\nu(x)\}$. Moreover, there exists $\delta^\nu > 0$ such that $(\cup_{k \notin K^\nu(x)} \operatorname{cl} R_k^\nu) \cap I\!B(x, \delta^\nu) = \emptyset$. This implies that $\lim_{\delta \searrow 0} \inf_{x' \in \boldsymbol{B}(x,\delta)} s^\nu(x') = s^\nu(x)$, which establishes that $s^\nu \in \text{e-spl}(\mathcal{R}^\nu)$.

We next show epi-convergence of s^ν to f. Let $x \in B$ be arbitrary. By the lsc of f, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $f(x') \geq f(x) - \varepsilon$ whenever $x' \in I\!B(x, \delta)$. Consequently,

$$\sigma_k^\nu \geq \min\left\{\nu, \inf_{\operatorname{cl} R_k^\nu} f\right\} \geq \min\{\nu, f(x) - \varepsilon\} \text{ when } R_k^\nu \subset I\!B(x, \delta).$$

Since $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ is an infinite refinement, there exists a $\bar{\nu}$ and $\delta_0 \in (0, \delta)$ such that $R_k^\nu \subset I\!B(x, \delta)$ for every $\nu \geq \bar{\nu}$ and k satisfying $R_k^\nu \cap I\!B(x, \delta_0) \neq \emptyset$. Consequently, for $x' \in B \cap I\!B(x, \delta_0/2)$ and $\nu \geq \bar{\nu}$,

$$s^\nu(x') = \min_{k=1,\ldots,N^\nu} \left\{\sigma_k^\nu \ : \ x' \in \operatorname{cl} R_k^\nu\right\} \geq \min\{\nu, f(x) - \varepsilon\}$$

because every k and $\nu \geq \bar{\nu}$ with $x' \in \operatorname{cl} R_k^\nu$ must also have $R_k^\nu \subset I\!B(x, \delta)$. For $x' \in I\!B(x, \delta_0/2) \setminus B$, $s^\nu(x') = \infty$. These fact establishes that, for every sequence $x^\nu \to x \in B$, $\operatorname{liminf} s^\nu(x^\nu) \geq \operatorname{liminf} \min\{\nu, f(x) - \varepsilon\} = f(x) - \varepsilon$, where the possibility that $f(x) = \infty$ is included. Since ε is arbitrary, $\operatorname{liminf} s^\nu(x^\nu) \geq f(x)$. This inequality also holds for $x \notin B$ also because B is closed and $s^\nu = f = \infty$ on $X \setminus B$. Next, again let $x \in B$ be arbitrary. By construction,

$$s^\nu(x) \leq \sigma_k^\nu \leq \max\left\{-\nu, \inf_{\operatorname{cl} R_k^\nu} f\right\} \text{ for every } k \text{ satisfying } x \in \operatorname{cl} R_k^\nu.$$

Hence, $s^\nu(x) \leq \max\{-\nu, f(x)\}$. Set $x^\nu = x$ for all ν and we obtain that $\operatorname{limsup} s^\nu(x^\nu) = \operatorname{limsup} s^\nu(x) \leq f(x)$. If $x \notin B$, then $f(x) = \infty$ and the previous inequality holds trivially for any x^ν. Thus, we have established epi-convergence of s^ν to f.

Since $\operatorname{liminf} s^\nu(x) \geq f(x)$ holds by virtue of the established epi-convergence, we also have that $s^\nu(x) \to f(x)$ for all $x \in X$, which establishes the pointwise convergence. The fact that $s^\nu(x) \leq \max\{-\nu, f(x)\}$ for all $x \in X$ is settled already. □

The proof of Theorem 5.5 is constructive. Given a partition $\{R_k\}_{k=1}^N$, an approximating epi-spline to a lsc function f is essentially the piecewise constant function given at $x \in R_k$ by $\inf_{R_k} f$.

We next examine approximation of functions in lsc-fcns(X) by epi-splines in the sense of the aw-distance, which requires us to adopt a "uniformity" requirement on infinite refinements. As we see below, this imposes further restrictions on the underlying space (X, d).

5.6 Definition (locally uniform infinite refinement) *A sequence $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ of partitions of a closed set $B \subseteq X$, with $\mathcal{R}^\nu = \{R_k^\nu\}_{k=1}^{N^\nu}$, is a locally uniform infinite refinement of B if*

for every $x \in B, \rho \geq 0,$ and $\varepsilon > 0,$ there exists a $\bar{\nu} \in I\!N$ such that
$R_k^\nu \subset I\!B(y, \varepsilon)$ for every $y \in I\!B(x, \rho) \cap B, \nu \geq \bar{\nu},$ and k satisfying $y \in \operatorname{cl} R_k^\nu$.

It should be apparent that a locally uniform infinite refinement is also an infinite refinement. A locally uniform infinite refinement needs to have $\bar{\nu}$ that applies not only at a single point x, as in the case of an infinite refinement, but for all points in arbitrarily large balls. Naturally, compactness ensures such a property as established next.

5.7 Proposition (sufficient condition for locally uniform infinite refinement) *There exists a locally uniform infinite refinement of every nonempty solid set $B \subseteq X$ for which $B \cap \mathbb{B}(x, r)$ is compact for all $x \in B$ and $r > 0$.*

Proof. Let $x^0 \in B$, $\nu \in \mathbb{N}$, and $B^\nu = \mathbb{B}(x^0, \nu) \cap B$. Since B^ν is compact, there exists $M^\nu < \infty$ and $\{x^{\nu,k}\}_{k=1}^{M^\nu} \subset B^\nu$ such that $\cup_{k=1}^{M^\nu} \mathbb{B}(x^{\nu,k}, 1/\nu) \supset B^\nu$. Set $R_1^\nu = \operatorname{int}(B^\nu \cap \mathbb{B}(x^{\nu,1}, 1/\nu))$ and recursively

$$R_k^\nu = \operatorname{int}\left(\left(B^\nu \cap \mathbb{B}(x^{\nu,k}, 1/\nu)\right) \setminus \cup_{l=1}^{k-1} R_l^\nu\right) \text{ for } k = 2, 3, ..., M^\nu.$$

Set $N^\nu = M^\nu + 1$ and let $R_{N^\nu}^\nu = \operatorname{int}(B \setminus \mathbb{B}(x^0, \nu))$. We observe that some R_k^ν maybe empty, but that is immaterial. Obviously, R_k^ν, $k = 1, ..., N^\nu$, are open and nonoverlapping, and $\cup_{k=1}^{N^\nu} R_k^\nu \subset B$. Since B is closed, we also have that $\cup_{k=1}^{N^\nu} \operatorname{cl} R_k^\nu \subset B$. We conclude that $\cup_{k=1}^{N^\nu} \operatorname{cl} R_k^\nu \supset B$ after following the same argument as in the proof of Proposition 5.4. Therefore, $\{R_k^\nu\}_{k=1}^{N^\nu}$ is a partition of B.

We next show that Definition 5.6 holds. Let $x \in B$, $\rho \geq 0$, and $\varepsilon > 0$. Set $\bar{\nu}$ equal to the smallest integer no smaller than $\max\{2/\varepsilon, d(x, x^0) + \varepsilon + \rho\}$. By construction, for every $\nu \geq \bar{\nu}$ and $k = 1, ..., N^\nu - 1$, $\sup\{d(y, y') \;:\; y, y' \in R_k^\nu\} \leq \varepsilon$. Moreover, for such ν, $\mathbb{B}(y, \varepsilon) \cap B \subset B^\nu$ for all $y \in \mathbb{B}(x, \rho) \cap B$ and therefore $R_k^\nu \subset \mathbb{B}(y, \varepsilon)$ whenever $y \in \operatorname{cl} R_k^\nu$ and $y \in \mathbb{B}(x, \rho) \cap B$. This establishes the result. □

Next we strengthen Theorem 5.5 by considering the aw-distance. We say that an epi-spline s on X is *rational* if $s(x)$ is a rational number for every $x \in \operatorname{dom} s$. The subset of rational epi-splines in e-spl($\mathcal{R}$) is denoted by r-spl($\mathcal{R}$).

5.8 Theorem (rational epi-splines dense in lsc functions) *If $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ is a locally uniform infinite refinement of a nonempty closed set $B \subseteq X$, then*

$$\bigcup_{\nu=1}^{\infty} \text{r-spl}(\mathcal{R}^\nu) \textit{ is dense in } (\textit{lsc-fcns}(B), \mathbb{d}).$$

Proof. Let $f \in$ lsc-fcns(B) and $\mathcal{R}^\nu = \{R_k^\nu\}_{k=1}^{N^\nu}$. For every $\nu \in \mathbb{N}$ and R_k^ν, $k = 1, 2, ..., N^\nu$, with $\inf_{\operatorname{cl} R_k^\nu} f$ finite, let q_k^ν be a rational number in $[\inf_{\operatorname{cl} R_k^\nu} f - 1/\nu, \inf_{\operatorname{cl} R_k^\nu} f]$. Let $\sigma_k^\nu = \min\{\nu, \max\{-\nu, q_k^\nu\}\}$. For $\nu \in \mathbb{N}$ and R_k^ν, $k = 1, 2, ..., N^\nu$, with $\inf_{\operatorname{cl} R_k^\nu} f = \infty$, set $\sigma_k^\nu = \nu$ and for $\inf_{\operatorname{cl} R_k^\nu} f = -\infty$ set $\sigma_k^\nu = -\nu$. We proceed by constructing $s^\nu : X \to \overline{\mathbb{R}}$ as follows:

$$s^\nu(x) = \min_{k=1,\dots,N^\nu} \left\{\sigma_k^\nu \;:\; x \in \operatorname{cl} R_k^\nu\right\} \text{ for } x \in B, \text{ and } s^\nu(x) = \infty \text{ for } x \notin B.$$

A replication of the arguments in the proof of Theorem 5.5 establishes that $s^\nu \in$ r-spl($\mathcal{R}^\nu$).

We next show that for any $\rho \geq 0$, $\hat{\mathbb{d}}_\rho^0(s^\nu, f) \to 0$ as $\nu \to \infty$. By construction, if $\inf_{\operatorname{cl} R_k^\nu} f$ is finite, then

$$s^\nu(x) \leq \sigma_k^\nu \leq \max\left\{-\nu, q_k^\nu\right\} \leq \max\left\{-\nu, \inf_{\operatorname{cl} R_k^\nu} f\right\} \leq \max\{-\nu, f(x)\} \text{ when } x \in \operatorname{cl} R_k^\nu.$$

If $\inf_{\operatorname{cl} R_k^\nu} f = \pm\infty$, then $s^\nu(x) \leq \nu$ for $x \in \operatorname{cl} R_k^\nu$. Since this holds for all $k = 1, ..., N^\nu$, $s^\nu(x) \leq \max\{-\nu, f(x)\}$ for $x \in X$. For $\nu \geq \rho$, $x \in X$, and $\eta = 0$,

$$\inf_{\mathbb{B}(x,\eta)} s^\nu = s^\nu(x) \leq \max\{f(x), -\nu\} + \eta \leq \max\{f(x), -\rho\} + \eta.$$

Therefore even $\eta = 0$ satisfies the second collection of constraints in the definition of $\hat{d\!l}^0_\rho(s^\nu, f)$. We now turn to the other constraints in the definition with the roles of s^ν and f reversed. Let $\varepsilon > 0$ and $z \in B$. Since $\{\mathcal{R}^\nu\}_{\nu=1}^\infty$ is a locally uniform infinite refinement, there exists a $\bar{\nu} \in I\!\!N$ such that $R_k^\nu \subset I\!\!B(x, \varepsilon)$ for every $x \in B \cap I\!\!B(z, d(\tilde{x}, z) + \rho)$, $\nu \geq \bar{\nu}$, and k satisfying $x \in \mathrm{cl}\, R_k^\nu$. Let $x \in \mathrm{lev}_\rho\, s^\nu \cap I\!\!B_\rho$ and $\nu > \max\{\rho, \bar{\nu}\}$. Since $d(x, z) \leq d(x, \tilde{x}) + d(\tilde{x}, z) \leq \rho + d(\tilde{x}, z)$, $x \in B \cap I\!\!B(z, d(\tilde{x}, z) + \rho)$. There exists a $k_x^\nu \in \{1, ..., N^\nu\}$ such that $x \in \mathrm{cl}\, R^\nu_{k_x^\nu}$ and $\inf_{\mathrm{cl}\, R^\nu_{k_x^\nu}} f \leq s^\nu(x) + 1/\nu$ due to the fact that $s^\nu(x) \leq \rho < \nu$. Since k_x^ν is one of possibly several k for which $R_k^\nu \subset I\!\!B(x, \varepsilon)$ holds, we obtain that for $\nu > \max\{\rho, \bar{\nu}, 1/\varepsilon\}$,

$$\inf_{I\!\!B(x,\varepsilon)} f \leq \inf_{\mathrm{cl}\, R^\nu_{k_x^\nu}} f < s^\nu(x) + \varepsilon \leq \max\{s^\nu(x), -\rho\} + \varepsilon.$$

Consequently, the first collection of constraints in the definition of $\hat{d\!l}^0_\rho(s^\nu, f)$ is satisfied with $\eta = \varepsilon$. We then have that $\hat{d\!l}^0_\rho(s^\nu, f) \leq \varepsilon$ for all $\nu > \max\{\rho, \bar{\nu}, 1/\varepsilon\}$. Since $\varepsilon > 0$ is arbitrary, we conclude that $\hat{d\!l}^0_\rho(s^\nu, f) \to 0$ as $\nu \to \infty$. By Propositions 3.1 and 3.2, this suffices to establish that $d\!l(s^\nu, f) \to 0$. $\square$

Next, we give a rate of convergence result for epi-splines, which is novel even for $I\!\!R^n$. The error in approximation of a lsc function by an epi-spline is bounded by the "size" of the open sets making up its partition. Specifically, for a partition $\mathcal{R} = \{R_k\}_{k=1}^N$ of a nonempty closed set $B \subseteq X$ and $\rho \geq 0$, we define its *meshsize* as

$$m_\rho(\mathcal{R}) := \inf \big\{\eta \geq 0 \ : \ R_k \subset I\!\!B(x, \eta) \text{ for all } x \in B \cap I\!\!B_\rho \text{ and } k \text{ satisfying } x \in \mathrm{cl}\, R_k\big\}.$$

5.9 Theorem (rate of convergence of epi-splines) *For a partition $\mathcal{R} = \{R_k\}_{k=1}^N$ of a nonempty closed set $B \subseteq X$ and $\rho \geq 0$, we have that*

$$\text{for every } f \in \text{lsc-fcns}(B), \text{ there exists an } s \in \text{e-spl}(\mathcal{R}) \text{ such that } \hat{d\!l}^0_\rho(s, f) \leq m_\rho(\mathcal{R}).$$

If $m_\rho(\mathcal{R}) > 0$, then s can be selected to be rational.

Proof. Let $f \in$ lsc-fcns(B). We start with the case of $m_\rho(\mathcal{R}) > 0$ and set $\bar{\nu} > \max\{m_\rho(\mathcal{R}), \rho\}$. For every R_k, $k = 1, 2, ..., N$, with $\inf_{\mathrm{cl}\, R_k} f$ finite, let q_k be a rational number in $[\inf_{\mathrm{cl}\, R_k} f - m_\rho(\mathcal{R}), \inf_{\mathrm{cl}\, R_k} f]$. Moreover, let $\sigma_k = \min\{\bar{\nu}, \max\{-\bar{\nu}, q_k\}\}$. For R_k, $k = 1, 2, ..., N$, with $\inf_{\mathrm{cl}\, R_k} f = \infty$, set $\sigma_k = \bar{\nu}$ and for $\inf_{\mathrm{cl}\, R_k} f = -\infty$ set $\sigma_k = -\bar{\nu}$. We proceed by constructing $s : X \to \overline{I\!\!R}$ as follows:

$$s(x) = \min_{k=1,...,N} \Big\{\sigma_k \ : \ x \in \mathrm{cl}\, R_k\Big\} \text{ for } x \in B, \text{ and } s(x) = \infty \text{ for } x \notin B.$$

By the same arguments as in the proof of Theorem 5.5, we conclude that $s \in$ r-spl$(\mathcal{R})$.

We next establish the approximation error associated with s. Mimicking the lines of reasoning in the proof of Theorem 5.8, we obtain that $\inf_{I\!\!B(x,\eta)} s \leq \max\{f(x), -\rho\} + \eta$ holds with $\eta = 0$ for all $x \in X$. Next we reverse the roles of s and f. Let $x \in \mathrm{lev}_\rho\, s \cap I\!\!B_\rho$. Certainly, all k with $x \in \mathrm{cl}\, R_k$ has $R_k \subset I\!\!B(x, m_\rho(\mathcal{R}))$. There exists a k_x such that $x \in \mathrm{cl}\, R_{k_x}$ and $\inf_{\mathrm{cl}\, R_{k_x}} f \leq s(x) + m_\rho(\mathcal{R})$ due to the fact that $s(x) \leq \rho < \bar{\nu}$. Since k_x is one of possibly several k for which $R_k \subset I\!\!B(x, m_\rho(\mathcal{R}))$ holds, we obtain that

$$\inf_{I\!\!B(x, m_\rho(\mathcal{R}))} f \leq \inf_{\mathrm{cl}\, R_{k_x}} f \leq s(x) + m_\rho(\mathcal{R}) \leq \max\{s(x), -\rho\} + m_\rho(\mathcal{R}).$$

We can therefore conclude that $\hat{d\!l}^0_\rho(s,f) \leq m_\rho(\mathcal{R})$ when $m_\rho(\mathcal{R}) > 0$. If $m_\rho(\mathcal{R}) = 0$, the same arguments hold except that q_k could be irrational and $s \in$ e-spl$(\mathcal{R})$. $\square$

An example illustrates Theorem 5.9. In optimization over X approximations might arise from approximations of X by a simpler set, say $\hat{X}$. As seen in Proposition 3.3, the error introduced by this approximation is largely given by $\hat{d\!l}_\rho(\hat{X}, X)$. To make this concrete, let $(X,d) = (\text{lsc-fcns}(I\!R^n), d\!l)$ and $\mathcal{R}$ be a partition of $I\!R^n$ that defines the simpler set $\hat{X} = \text{e-spl}(\mathcal{R})$. We recall from Propositions 3.1 and 3.2 (applied with $X = I\!R^n$ and $\tilde{x} = 0$) that for $f \in \text{lsc-fcns}(I\!R^n)$ and $s \in \text{e-spl}(\mathcal{R})$,

$$d\!l(s,f) \leq (1 - e^{-\gamma})\hat{d\!l}^0_{\gamma'}(s,f) + e^{-\gamma}[\max\{\delta_s, \delta_f\} + \gamma + 1] \text{ for any } \gamma \geq 0,$$

where $\delta_s = \text{dist}(0, \text{epi}\, s)$, $\delta_f = \text{dist}(0, \text{epi}\, f)$, and $\gamma' \geq 2\gamma + \max\{\delta_s, \delta_f\}$. For every $f \in \text{lsc-fcns}(I\!R^n)$ there exists by Theorem 5.9 an $s_f \in \text{e-spl}(\mathcal{R})$ such that $\hat{d\!l}^0_{\gamma'}(s_f, f) \leq m_{\gamma'}(\mathcal{R})$. Moreover, there exists a $\delta > \sup\{\max\{\delta_{s_f}, \delta_f\} \;:\; f \in \text{lsc-fcns}(I\!R^n) \cap I\!B_\rho\}$. For every $\gamma \geq 0$ and $f \in \text{lsc-fcns}(I\!R^n) \cap I\!B_\rho$, we have then

$$\text{dist}(f, \text{e-spl}(\mathcal{R})) \leq d\!l(s_f, f) \leq (1 - e^{-\gamma}) m_{2\gamma+\delta}(\mathcal{R}) + e^{-\gamma}[\delta + \gamma + 1].$$

Thus,

$$\hat{d\!l}_\rho(\text{e-spl}(\mathcal{R}), \text{lsc-fcns}(I\!R^n)) \leq (1 - e^{-\gamma}) m_{2\gamma+\delta}(\mathcal{R}) + e^{-\gamma}[\delta + \gamma + 1].$$

We end the section with an observation that the existence of a locally uniform infinite refinement is intimately tied to compactness of balls; Proposition 5.7 shows that such compactness is a sufficient conditions. The fact that it is also necessary when B is complete is stated next.

5.10 Proposition (necessary condition for locally uniform infinite refinement) *If there exists a locally uniform infinite refinement of a nonempty closed set $B \subseteq X$ and B is complete, then every closed ball in X must have a compact intersection with B.*

Proof. By Theorem 5.8, the assumption implies that $(\text{lsc-fcns}(B), d\!l)$ is separable. In view of [10, Theorem 3.1.4] we know that this takes place if and only if closed and bounded subsets of (B,d) are totally bounded. Thus, closed balls in X intersected with B are totally bounded. Their compactness then follows from the assumption of completeness. $\square$

Appendix

Proof of Proposition 3.1. Item (i) follows immediately from the definitions. For (ii), first observe that for a nonempty set $C \subset X \times I\!R$ and $\bar{x}, \bar{y} \in X \times I\!R$ we have for any $\bar{z} \in C$, $\text{dist}(\bar{x}, C) \leq \bar{d}(\bar{x}, \bar{z}) \leq \bar{d}(\bar{x}, \bar{y}) + \bar{d}(\bar{y}, \bar{z})$. Minimizing over $\bar{z} \in C$, we obtain that $\text{dist}(\bar{x}, C) \leq \bar{d}(\bar{x}, \bar{y}) + \text{dist}(\bar{y}, C)$ and thus

$$|\,\text{dist}(\bar{x}, C) - \text{dist}(\bar{y}, C)| \leq \bar{d}(\bar{x}, \bar{y}).$$

Second, in the notation $h(\bar{x}) = |\,\text{dist}(\bar{x}, C) - \text{dist}(\bar{x}, D)|$ for nonempty $C, D \subset X \times I\!R$, we have that

$$\begin{aligned} |h(\bar{x}) - h(\bar{y})| &\leq |[\text{dist}(\bar{x}, C) - \text{dist}(\bar{x}, D)] - [\text{dist}(\bar{y}, C) - \text{dist}(\bar{y}, D)]| \\ &\leq |\,\text{dist}(\bar{x}, C) - \text{dist}(\bar{y}, C)| + |\,\text{dist}(\bar{x}, D) - \text{dist}(\bar{y}, D)| \leq 2\bar{d}(\bar{x}, \bar{y}). \end{aligned}$$

Let $\varepsilon > 0$. For every $\bar{y} \in \mathbb{S}_{\rho'}$, there exists an $\bar{x} \in \mathbb{S}_\rho$ such that $\bar{d}(\bar{x}, \bar{y}) \leq e(\mathbb{S}_{\rho'}, \mathbb{S}_\rho) + \varepsilon$. Combining these last two facts and replacing C and D by epi-graphs, we obtain that

$$\begin{aligned} d\!\!l_{\rho'}(f,g) &= \sup\{|\operatorname{dist}(\bar{y}, \operatorname{epi} f) - \operatorname{dist}(\bar{y}, \operatorname{epi} g)| \ : \ \bar{y} \in \mathbb{S}_{\rho'}\} \\ &\leq \sup\{|\operatorname{dist}(\bar{x}, \operatorname{epi} f) - \operatorname{dist}(\bar{x}, \operatorname{epi} g)| + 2e(\mathbb{S}_{\rho'}, \mathbb{S}_\rho) + \varepsilon \ : \ \bar{x} \in \mathbb{S}_\rho\} = d\!\!l_\rho(f,g) + 2e(\mathbb{S}_{\rho'}, \mathbb{S}_\rho) + \varepsilon, \end{aligned}$$

which establishes (ii) after realizing that $\varepsilon > 0$ is arbitrary.

Next consider (iii). Suppose that $C, D \subset X \times I\!R$ are nonempty closed, $\varepsilon > 0$, and $\rho \geq 0$. We first show that

$$\operatorname{dist}(\cdot, D) \leq \operatorname{dist}(\cdot, C) + \varepsilon \text{ on } \mathbb{S}_\rho \text{ implies that } C \cap \mathbb{S}_\rho \subset D_\varepsilon^+ := \{\bar{x} \in X \times I\!R \ : \ \operatorname{dist}(\bar{x}, D) \leq \varepsilon\}.$$

The claim is trivial if $C \cap \mathbb{S}_\rho$ is empty. For nonempty $C \cap \mathbb{S}_\rho$, we have for every $\bar{x} \in C \cap \mathbb{S}_\rho$ that $\operatorname{dist}(\bar{x}, D) \leq \varepsilon$ and the implication follows. The translation of this fact to the context of epi-graphs establishes the lower bound in (iii). Second, we establish that

$$C \cap \mathbb{S}_{\rho'} \subset D_\varepsilon^+ \text{ implies } \operatorname{dist}(\cdot, D) \leq \operatorname{dist}(\cdot, C) + \varepsilon \text{ on } \mathbb{S}_\rho \text{ for } \rho' > 2\rho + \operatorname{dist}((\tilde{x}, 0), C). \tag{8}$$

For $\bar{z} \in C \cap \mathbb{S}_{\rho'} \subset D_\varepsilon^+$ and $\bar{x} \in X \times I\!R$, $\operatorname{dist}(\bar{x}, D) \leq \bar{d}(\bar{x}, \bar{z}) + \operatorname{dist}(\bar{z}, D) \leq \bar{d}(\bar{x}, \bar{z}) + \varepsilon$. The minimization over $\bar{z} \in C \cap \mathbb{S}_{\rho'}$ gives that

$$\operatorname{dist}(\bar{x}, D) \leq \operatorname{dist}(\bar{x}, C \cap \mathbb{S}_{\rho'}) + \varepsilon. \tag{9}$$

This holds trivially if $C \cap \mathbb{S}_{\rho'} = \emptyset$. Suppose that $\bar{x} \in \mathbb{S}_\rho$ and $\rho' > 2\rho + \operatorname{dist}((\tilde{x}, 0), C)$. For every ν, there exists $\bar{y}^\nu \in C$ such that $\bar{d}(\bar{x}, \bar{y}^\nu) \leq \operatorname{dist}(\bar{x}, C) + 1/\nu$. Moreover,

$$\begin{aligned} \bar{d}((\tilde{x}, 0), \bar{y}^\nu) &\leq \bar{d}((\tilde{x}, 0), \bar{x}) + \bar{d}(\bar{x}, \bar{y}^\nu) = \bar{d}((\tilde{x}, 0), \bar{x}) + \operatorname{dist}(\bar{x}, C) + 1/\nu \\ &\leq \bar{d}((\tilde{x}, 0), \bar{x}) + \bar{d}(\bar{x}, (\tilde{x}, 0)) + \operatorname{dist}((\tilde{x}, 0), C) + 1/\nu \leq \rho + \rho + \operatorname{dist}((\tilde{x}, 0), C) + 1/\nu. \end{aligned}$$

Since $\rho' > 2\rho + \operatorname{dist}((\tilde{x}, 0), C)$, there exists a $\bar{\nu}$ such that $\bar{y}^\nu \in C \cap \mathbb{S}_{\rho'}$ for all $\nu \geq \bar{\nu}$. For such ν,

$$\operatorname{dist}(\bar{x}, C \cap \mathbb{S}_{\rho'}) \leq \bar{d}(\bar{x}, \bar{y}^\nu) \leq \operatorname{dist}(\bar{x}, C) + 1/\nu.$$

Letting $\nu \to \infty$ in this expression and observing that $\operatorname{dist}(\bar{x}, C \cap \mathbb{S}_{\rho'}) \geq \operatorname{dist}(\bar{x}, C)$ generally, we obtain that $\operatorname{dist}(\bar{x}, C \cap \mathbb{S}_{\rho'}) = \operatorname{dist}(\bar{x}, C)$, which together with (9) establishes (8). The implication in (8) directly confirms the upper bound in (iii). If (X, d) is proper, then in view of Lemma 2.4 we can take $\bar{y}^\nu$ above to satisfy $\bar{d}(\bar{x}, \bar{y}^\nu) = \operatorname{dist}(\bar{x}, C)$ for all ν. Thus, the need for the $1/\nu$ term vanishes and the stronger statement given at the end of the theorem is established.

Item (iv) follows trivially from the definition of $d\!\!l_\rho$. For (v) and (vi), we follow the lines of arguments in the proof of [24, Lemma 4.41]. Clearly,

$$d\!\!l(f,g) = \int_0^\rho d\!\!l_\tau(f,g) e^{-\tau} d\tau + \int_\rho^\infty d\!\!l_\tau(f,g) e^{-\tau} d\tau.$$

Since $d\!\!l_\tau(f,g)$ is nondecreasing as τ increases, we have that

$$d\!\!l_0(f,g) \int_0^\rho e^{-\tau} d\tau \leq \int_0^\rho d\!\!l_\tau(f,g) e^{-\tau} d\tau \leq d\!\!l_\rho(f,g) \int_0^\rho e^{-\tau} d\tau.$$

and

$$\mathbb{d}\!l_\rho(f,g)\int_\rho^\infty e^{-\tau}d\tau \le \int_\rho^\infty \mathbb{d}\!l_\tau(f,g)e^{-\tau}d\tau \le \int_\rho^\infty [\max\{\delta_f,\delta_g\}+\tau]e^{-\tau}d\tau,$$

where the last inequality follows (iv). Carrying out the integrations on the left- and right-hand sides, we obtain (v) and (vi).

The lower bound in (vii) is obtained by letting ρ tend to infinity in (v). Item (vi) with $\rho = 0$ furnishes the upper bound. ◻

Proof of Proposition 3.2. We start by establishing the finiteness of $\hat{\mathbb{d}\!l}{}^0_\rho(f,g)$ for any $\rho \ge 0$. Since $f,g \in$ lsc-fcns(X), there exist $\bar{x},\bar{y} \in X$ such that $g(\bar{x}), f(\bar{y}) < \infty$. Let $\bar{\eta} = \rho + \max\{d(\tilde{x},\bar{x}), d(\tilde{x},\bar{y}), g(\bar{x}), f(\bar{y})\}$, which is finite. Then, for $x \in \mathrm{lev}_\rho f \cap \mathbb{B}_\rho$, $d(x,\bar{x}) \le d(x,\tilde{x}) + d(\tilde{x},\bar{x}) \le \rho + (\bar{\eta}-\rho) = \bar{\eta}$ and thus $\inf_{\mathbb{B}(x,\bar{\eta})} g \le g(\bar{x}) \le \bar{\eta} - \rho \le \max\{f(x),-\rho\} + \bar{\eta}$. Likewise, for $x \in \mathrm{lev}_\rho g \cap \mathbb{B}_\rho$, $\inf_{\mathbb{B}(x,\bar{\eta})} f \le \max\{g(x),-\rho\} + \bar{\eta}$. Thus, $\hat{\mathbb{d}\!l}{}^0_\rho(f,g) \le \bar{\eta}$, which establishes the rightmost inequality.

Obviously, $\hat{\mathbb{d}\!l}_\rho(f,g) = \inf\{\eta \ge 0 \;:\; \hat{\mathbb{d}\!l}_\rho(f,g) \le \eta\}$, which in view of the minimization taking place in the definition of $\hat{\mathbb{d}\!l}{}^\delta_\rho(f,g)$ motivates the examination of the relation $\hat{\mathbb{d}\!l}_\rho(f,g) \le \eta$ for $\eta \ge 0$. It follows directly from the definition of $\hat{\mathbb{d}\!l}_\rho$ that

$$\hat{\mathbb{d}\!l}_\rho(f,g) \le \eta \text{ if and only if } \mathrm{dist}(\bar{x}, \mathrm{epi}\, g) \le \eta \;\forall \bar{x} \in \mathrm{epi}\, f \cap \mathbb{S}_\rho \text{ and } \mathrm{dist}(\bar{x}, \mathrm{epi}\, f) \le \eta \;\forall \bar{x} \in \mathrm{epi}\, g \cap \mathbb{S}_\rho,$$

which in turn is equivalent to having

$$\big(\mathrm{epi}\, f\big) \cap \mathbb{S}_\rho \subset D^+_\eta(g) \text{ and } \big(\mathrm{epi}\, g\big) \cap \mathbb{S}_\rho \subset D^+_\eta(f),$$

where $D^+_\eta(f) := \{\bar{x} \in X \times \mathbb{R} \;:\; \mathrm{dist}(\bar{x}, \mathrm{epi}\, f) \le \eta\}$ and similarly for $D^+_\eta(g)$. By virtue of being defined in terms of distance to an epigraph, we have that $(x,y_0) \in D^+_\eta(g)$ implies $(x,x_0) \in D^+_\eta(g)$ for all $x_0 \ge y_0$. Thus,

$$\big(\mathrm{epi}\, f\big) \cap \mathbb{S}_\rho \subset D^+_\eta(g) \text{ if and only if } (x, f_\rho(x)) \in D^+_\eta(g) \text{ for all } x \in \mathrm{dom}\, f_\rho,$$

where the function $f_\rho : X \to \overline{\mathbb{R}}$ is given by $f_\rho(x) = \max\{f(x), -\rho\}$ if $x \in \mathrm{lev}_\rho f \cap \mathbb{B}_\rho$ and $f_\rho(x) = \infty$ otherwise. By definition, a point $(x,x_0) \in D^+_\eta(g)$ if and only if $\mathrm{dist}((x,x_0), \mathrm{epi}\, g) \le \eta$. The latter condition is more explicitly stated as

$$\inf\Big\{\max\{d(x,y), |x_0 - y_0|\} \;:\; g(y) \le y_0, y \in X, y_0 \in \mathbb{R}\Big\} \le \eta.$$

We are now in a position to establish the lower bound and let $\delta > 0$. Collecting the above facts, we find that if $\hat{\mathbb{d}\!l}_\rho(f,g) \le \eta$, then

$$\inf\Big\{\max\{d(x,y), |f_\rho(x) - y_0|\} \;:\; g(y) \le y_0, y \in X, y_0 \in \mathbb{R}\Big\} \le \eta \text{ for } x \in \mathrm{dom}\, f_\rho.$$

Let $x \in \mathrm{dom}\, f_\rho$. Hence, for every $\varepsilon \in (0,\delta]$, there exists $(y_\varepsilon, y_{0\varepsilon}) \in X \times R$ such that $g(y_\varepsilon) \le y_{0\varepsilon}$, $d(x,y_\varepsilon) \le \eta + \varepsilon$, and $|f_\rho(x) - y_{0\varepsilon}| \le \eta + \varepsilon$. Consequently, $g(y_\varepsilon) \le f_\rho(x) + \eta + \varepsilon$ and $y_\varepsilon \in \mathbb{B}(x, \eta+\varepsilon)$. Moreover, $\inf_{\mathbb{B}(x,\eta+\delta)} g \le \inf_{\mathbb{B}(x,\eta+\varepsilon)} g \le f_\rho(x) + \eta + \varepsilon$. Since this relation holds for all $\varepsilon \in (0,\delta]$, $\inf_{\mathbb{B}(x,\eta+\delta)} g \le f_\rho(x) + \eta$ for $x \in \mathrm{dom}\, f_\rho$. A parallel development gives identical results with the roles of f and g reversed, where we let $g_\rho(x) = \max\{g(x), -\rho\}$ if $x \in \mathrm{lev}_\rho\, g \cap I\!\!B_\rho$ and $g_\rho(x) = \infty$ otherwise. Specifically, we have that $\inf_{B(x,\eta+\delta)} f \le g_\rho(x) + \eta$ for $x \in \mathrm{dom}\, g_\rho$. Since $x \in \mathrm{dom}\, f_\rho$ if and only if $x \in \mathrm{lev}_\rho\, f \cap I\!\!B_\rho$, we also have that

$$\inf_{B(x,\eta+\delta)} g \le f_\rho(x) + \eta \text{ for } x \in \mathrm{lev}_\rho\, f \cap I\!\!B_\rho.$$

and similarly

$$\inf_{B(x,\eta+\delta)} f \le g_\rho(x) + \eta \text{ for } x \in \mathrm{lev}_\rho\, g \cap I\!\!B_\rho.$$

The lower bound then follows after observing that these relations hold, in particular, for $\eta = \hat{d\!l}_\rho(f, g)$.

Next, we address the upper bound. Suppose that $\eta \ge 0$ satisfies

$$\inf_{B(x,\eta)} g \le \max\{f(x), -\rho\} + \eta, \forall x \in \mathrm{lev}_\rho\, f \cap I\!\!B_\rho.$$

As above, this means that $\inf_{B(x,\eta)} g \le f_\rho(x) + \eta$ for $x \in \mathrm{dom}\, f_\rho$. We now examine this relation for a fixed $x \in \mathrm{dom}\, f_\rho$. For every $\varepsilon > 0$, there exists a $y_\varepsilon \in X$ such that $d(x, y_\varepsilon) \le \eta$ and $g(y_\varepsilon) \le f_\rho + \eta + \varepsilon$. Set $y_{0\varepsilon} = \max\{g(y_\varepsilon), f_\rho(x) - \eta - \varepsilon\}$. Thus, $g(y_\varepsilon) \le y_{0\varepsilon}$ and

$$f_\rho(x) - y_{0\varepsilon} \le f_\rho(x) - (f_\rho(x) - \eta - \varepsilon) = \eta + \varepsilon.$$

Moreover, $|f_\rho(x) - y_{0\varepsilon}| \le \eta + \varepsilon$. We have therefore established that

$$\max\{d(x, y_\varepsilon), |f_\rho(x) - y_{0\varepsilon}|\} \le \eta + \varepsilon \text{ and } g(y_\varepsilon) \le y_{0\varepsilon}.$$

Since this holds for all $\varepsilon > 0$,

$$\inf \Big\{ \max\{d(x, y), |f_\rho(x) - y_0|\} \; : \; g(y) \le y_0, y \in X, y_0 \in I\!\!R \Big\} \le \eta \text{ for } x \in \mathrm{dom}\, f_\rho.$$

Equivalently, $(x, f_\rho(x)) \in D^+_\eta(g)$ for $x \in \mathrm{dom}\, f_\rho$. A parallel development with the roles of f and g reversed, leads to $(x, g_\rho(x)) \in D^+_\eta(f)$ for $x \in \mathrm{dom}\, g_\rho$. The implications established in the beginning of the proof show that we then must have that $\hat{d\!l}_\rho(f, g) \le \eta$. In view of the definition of $\hat{d\!l}^0_\rho(f, g)$, it is possible to repeat the above arguments with η replaced by η^ν and have $\eta^\nu \searrow \hat{d\!l}^0_\rho(f, g)$ as well as $\hat{d\!l}_\rho(f, g) \le \eta^\nu$. This established the upper bound of the theorem.

We next consider the last assertion under the additional assumption that the space is proper. Again, suppose that $\hat{d\!l}_\rho(f, g) \le \eta$ and, thus,

$$\mathrm{dist}((x, f_\rho(x)), \mathrm{epi}\, g) \le \eta \text{ for } x \in \mathrm{dom}\, f_\rho \text{ and } \mathrm{dist}((x, g_\rho(x)), \mathrm{epi}\, f) \le \eta \text{ for } x \in \mathrm{dom}\, g_\rho.$$

Fix $x \in \mathrm{dom}\, f_\rho$. By Lemma 2.4 and the fact that $\mathrm{epi}\, g$ is a nonempty closed set, there exists $(y^*, y^*_0) \in X \times R$, with $g(y^*) \le y^*_0$, such that $\eta \ge \mathrm{dist}((x, f_\rho(x)), \mathrm{epi}\, g) = \bar{d}((x, f_\rho(x)), (y^*, y^*_0))$. Hence, $d(x, y^*) \le \eta$ and $|f_\rho(x) - y^*_0| \le \eta$, which leads to $g(y^*) \le f_\rho(x) + \eta$ and $y^* \in I\!\!B(x, \eta)$. This fact and a parallel development with the roles of g and f reversed give that

$$\inf_{B(x,\eta)} g \le f_\rho(x) + \eta \text{ for } x \in \mathrm{dom}\, f_\rho \text{ and } \inf_{B(x,\eta)} f \le g_\rho(x) + \eta \text{ for } x \in \mathrm{dom}\, g_\rho.$$

Repeating the last lines of reasoning that lead to the lower bound on $\hat{d\!l}_\rho(f,g)$, we conclude that under the additional assumption, the lower bound can be improved to $\hat{d\!l}^0_\rho(f,g)$. □

Proof of Proposition 3.3. We observe that $\hat{d\!l}_\rho(C,D) < \infty$ since C, D are nonempty. Let $\rho \in [0,\infty)$, $\rho' \in (\rho + \hat{d\!l}_\rho(C,D), \infty)$, and $\varepsilon \in (0, \rho' - \rho - \hat{d\!l}_\rho(C,D)]$. Set

$$\eta_\varepsilon = \sup_{A_\rho} |f-g| + \max\{1, \kappa(\rho')\} \left[\hat{d\!l}_\rho(C,D) + \varepsilon\right].$$

First, we establish that

$$\inf_{B(x,\eta_\varepsilon)}\{g + \iota_D\} \leq \max\{f(x) + \iota_C(x), -\rho\} + \eta_\varepsilon \text{ for } x \in \mathrm{lev}_\rho\{f + \iota_C\} \cap I\!B_\rho.$$

Suppose that $x \in \mathrm{lev}_\rho\{f + \iota_C\} \cap I\!B_\rho$, which of course implies that $x \in C$. There exists a $y \in D$ such that $d(x,y) \leq \inf\{d(x,y') : y' \in D\} + \varepsilon$. Thus, $\hat{d\!l}_\rho(C,D) \geq e(C \cap I\!B_\rho, D) \geq \mathrm{dist}(x,D) \geq d(x,y) - \varepsilon$ and $d(x,y) \leq \eta_\varepsilon$. Moreover, $d(\tilde{x},y) \leq d(\tilde{x},x) + d(x,y) \leq \rho + \hat{d\!l}_\rho(C,D) + \varepsilon \leq \rho'$. These facts and the Lipschitz continuity of g on $I\!B_{\rho'}$ imply that

$$\begin{aligned}
\inf_{B(x,\eta_\varepsilon)}\{g + \iota_D\} &\leq g(y) + \iota_D(y) = g(y) = g(y) - g(x) + g(x) - f(x) + f(x) \\
&\leq \kappa(\rho')d(x,y) + \sup_{A_\rho}|f-g| + \max\{f(x) + \iota_C(x), -\rho\} \\
&\leq \max\{f(x) + \iota_C(x), -\rho\} + \eta_\varepsilon,
\end{aligned}$$

which establishes the first claim. Second, following a parallel argument, we realize that

$$\inf_{B(x,\eta_\varepsilon)}\{f + \iota_C\} \leq \max\{g(x) + \iota_D(x), -\rho\} + \eta_\varepsilon \text{ for } x \in \mathrm{lev}_\rho\{g + \iota_D\} \cap I\!B_\rho.$$

Consequently, $\hat{d\!l}^0_\rho(f,g) \leq \eta_\varepsilon$. Since this holds for arbitrarily small $\varepsilon > 0$, the main conclusion follows. If (X,d) is proper, then the minimum distance between a point and a nonempty closed set is attained and the above arguments hold with $\varepsilon = 0$; see Lemma 2.4. This establishes that $\rho' = \rho + \hat{d\!l}_\rho(C,D)$ is permitted in this case. □

References

[1] H. Attouch. *Variational Convergence for Functions and Operators.* Applicable Mathematics Sciences. Pitman, 1984.

[2] H. Attouch, R. Lucchetti, and R. J-B Wets. The topology of the ρ-Hausdorff distance. *Annali di Matematica pura ed applicata*, CLX:303–320, 1991.

[3] H. Attouch and R. J-B Wets. A convergence theory for saddle functions. *Transactions of the American Mathematical Society*, 280(1):1–41, 1983.

[4] H. Attouch and R. J-B Wets. Isometries for the Legendre-Fenchel transform. *Transactions of the American Mathematical Society*, 296:33–60, 1986.

[5] H. Attouch and R. J-B Wets. Quantitative stability of variational systems: I. The epigraphical distance. *Transactions of the American Mathematical Society*, 328(2):695–729, 1991.

[6] H. Attouch and R. J-B Wets. Quantitative stability of variational systems: II. A framework for nonlinear conditioning. *SIAM J. Optimization*, 3:359–381, 1993.

[7] H. Attouch and R. J-B Wets. Quantitative stability of variational systems: III. ε-approximate solutions. *Mathematical Programming*, 61:197–214, 1993.

[8] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis.* Birkhäuser, 1990.

[9] D. Aze and J.-P. Penot. Recent quantitative results about the convergence of convex sets and functions. In *Functional Analysis and Approximations. Proceedings of the International Conference Bagni di Lucca*, pages 90–110. Pitagora Editrice, 1990.

[10] G. Beer. *Topologies on Closed and Closed Convex Sets*, volume 268 of *Mathematics and its Applications.* Kluwer, 1992.

[11] G. Beer and R. Lucchetti. Convex optimization and the epi-distance topology. *Transactions of the American Mathematical Society*, 327(2):795–813, 1991.

[12] G. Beer and R. Lucchetti. The epi-distance topology: continuity and stability with application to convex optimization. *Mathematics of Operations Research*, 17:715–726, 1992.

[13] J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems.* Springer, 2000.

[14] O. Bousquet and A. Elisseeff. Stability and generalization. *Journal of Machine Learning Research*, 2:499–526, 2002.

[15] D. Drusvyatskiy and A. S. Lewis. Tilt stability, uniform quadratic growth, and strong metric regularity of the subdifferential. *SIAM J. Optimization*, 23:256–267, 2013.

[16] A. C. Eberhard and R. Wenczel. A study of tilt-stable optimality and sufficient conditions. *Nonlinear Analysis*, 75:1260–1281, 2012.

[17] P. J. Laurent. *Approximation et optimisation.* Hermann, 1972.

[18] A. S. Lewis and S. Zhang. Partial smoothness, tilt stability, and generalized hessians. *SIAM J. Optimization*, 23:74–94, 2013.

[19] B. Mordukhovich. *Variational Analysis and Generalized Differentiation I: Basic Theory.* Grundlehren der mathematischen Wissenschaften. Springer, 2 edition, 2013.

[20] B. Mordukhovich. *Variational Analysis and Generalized Differentiation, II: Applications.* Grundlehren der mathematischen Wissenschaften. Springer, 2 edition, 2013.

[21] B. S. Mordukhovich, R. T. Rockafellar, and M. E. Sarabi. Characterizations of full stability in constrained optimization. *SIAM J. Optimization*, 23:1810–1849, 2013.

[22] U. Mosco. Convergence of convex sets and of solutions of variational inequalities. *Advances in Mathematics*, 3:510–585, 1969.

[23] E. Polak. *Optimization. Algorithms and Consistent Approximations*, volume 124 of *Applied Mathematical Sciences*. Springer, 1997.

[24] R.T. Rockafellar and R. J-B Wets. *Variational Analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaft*. Springer, 3rd printing-2009 edition, 1998.

[25] W. Römish and R. J-B Wets. Stability of ε-approximate solutions to convex stochastic programs. *SIAM J. Optimization*, 18(3):961–979, 2007.

[26] J. O. Royset and R. J-B Wets. From data to assessments and decisions: Epi-spline technology. In A. Newman, editor, *INFORMS Tutorials*. INFORMS, Catonsville, 2014.

[27] J. O. Royset and R. J-B Wets. Multivariate epi-splines and evolving function identification problems. *Set-Valued and Variational Analysis*, to appear, 2016.

[28] J. O. Royset and R. J-B Wets. Optimality functions and lopsided convergence. *J. Optimization Theory and Applications*, to appear, 2016.

[29] J. O. Royset and R. J-B Wets. Variational theory for optimization under stochastic ambiguity. In review, 2016.

[30] J.O. Royset and R. J-B Wets. Fusion of hard and soft information in nonparametric density estimation. *European J. of Operational Research*, 247(2):532–547, 2015.

[31] R. J-B Wets. Convergence of convex functions, variational inequalities, and convex optimization problems. In R. Cottle, R. Giannessi, and J.-L. Lions, editors, *Varitional Inequalities and Complementary Problems*, pages 375–403. Wiley, New York, 1980.

[32] R. A. Wijsman. Convergence of sequences of convex sets, cones and functions. *Bulletin of the American Mathematical Society*, 70:186–188, 1964.

[33] R. A. Wijsman. Convergence of sequences of convex sets, cones and functions. II. *Transactions of the American Mathematical Society*, 123:32–45, 1966.